

DOCUMENT RESUME

ED 037 771

CG 005 153

AUTHOR Adkins, Dorothy C.; And Others
TITLE Motivation to Achieve in School. Final Report.
INSTITUTION Hawaii Univ., Honolulu. Educational Research and
Development Center.
PUB DATE Jan 70
NOTE 170p.

EDRS PRICE EDRS Price MF-\$0.75 HC-\$8.60
DESCRIPTORS Academic Aspiration, *Achievement, Behavior,
*Curriculum Design, Educational Research, *Learning
Motivation, *Motivation, *Motivation Techniques,
Preschool Children, Research Projects

ABSTRACT

This research on motivation has been focused on (a) development of a measure of motivation of young children to achieve in school, a measure consonant with a five-faceted theory of constituents of motivation, and (b) a curriculum for teaching motivation to preschool children, also grounded in the theory. The resulting measure, Jumpgookies, is an objective-projective device consisting of dichotomous items. The test has been given to 1,607 children in the United States and 668 children in grades one, two, and four in Hawaii. Item difficulty indices and item test correlation coefficients have been studied, and data for each group and for various combinations of groups have been factor analyzed. Results of the first attempts to teach motivation to preschool children are regarded as sufficiently promising to warrant intensive efforts. Plans for extending this work to children both younger and older than preschoolers are being pursued. Follow-up of the groups involved in the study is also projected. (author/KJ)

F-OEO

ED037771

MOTIVATION TO ACHIEVE IN SCHOOL

CG 0 05153

**Education Research and Development Center
University of Hawaii**

ED037771

F-OEO
B89-4576

University of Hawaii
Head Start Research Center
Dorothy C. Adkins, Director

Final Report

Motivation To Achieve in School

Dorothy C. Adkins, Professor and Researcher
Bonnie L. Ballif, Assistant Professor, Assistant Researcher,
and Consultant (now at Fordham University)

Assisted by:

Vidya Bhushan, Assistant Professor and Assistant Researcher
Renato Espinosa, Assistant Professor and Assistant Researcher
Patricia Nash, Assistant in Research
Nancy Johngren, Assistant in Research
John Gilheney, Assistant in Research
Diana Selser, Assistant in Research

The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D. C. 20506. The opinions expressed herein are those of the authors and should not be construed as representing the opinions or policy of any agency of the United States Government.

Contract No. OEO B89-4576
Contract No. OEO 4121

Education Research and Development Center
David G. Ryans, Director
College of Education
University of Hawaii
Honolulu, Hawaii

January 1970

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

TABLE OF CONTENTS

List of Tables	i
Foreword	iv
Abstract	vii
Chapter 1--Problems Investigated	1
Chapter 2--Theoretical Orientation	6
Chapter 3--Early Methodological Considerations in the Development of <u>Gumpgookies</u>	16
Chapter 4--Basic Test and Item Statistics for <u>Gumpgookies</u>	29
Chapter 5--Data Pertaining to the Validity and Reliability of <u>Gumpgookies</u>	54
Chapter 6--Factor Analyses of Randomized Forms of <u>Gumpgookies</u> for First-, Second-, and Fourth-Graders and for Ten Groups of Preschoolers	75
Chapter 7--Explorations in Teaching Preschoolers Motivation To Achieve	120
Chapter 8--Summary and Proposed Future Research	138
Appendix	150
References	154

LIST OF TABLES

CHAPTER 4

Table 4.1--Distribution Statistics for the 100-Item Group Form of <u>Gumpgookies</u> for Three Grades	30
Table 4.2--Difficulty Values (Percentages Passing, D) and Item-Test Point-Biserial Correlation Coefficients (C) for the Items in the 100-Item Group Form of <u>Gumpgookies</u> (First Grade N = 212; Second Grade N = 223; Fourth Grade N = 233)	31
Table 4.3--Means, Standard Deviations, and Other Statistics for the 75-Item <u>Gumpgookies</u> Test, Ten Subsamples . . .	39
Table 4.4--Means and Standard Deviations of Age and <u>Gumpgookies</u> Score and the Correlation Between Them for Males in Ten Subsamples and the Total Group of Males	40
Table 4.5--Means and Standard Deviations of Age and <u>Gumpgookies</u> Score and the Correlation Between Them for Females in Ten Subsamples and the Total Group of Females . . .	41
Table 4.6--Difficulty Values (Percentages Passing) and Item-Test Correlations (Point-Biserials) of Items in the 75-Item <u>Gumpgookies</u> for Ten Subsamples	43
Table 4.7--Correlations (Pearson r 's) Among Difficulty Indices of <u>Gumpgookies</u> Items for Ten Subsamples . . .	48
Table 4.8--Correlations (Pearson r 's) Among <u>Gumpgookies</u> Item-Test Correlation Coefficients for Ten Subsamples	49
Table 4.9--Orthogonally Transformed Factor Loadings Based Upon Factor Analysis of the Item Difficulty Values for Ten Subsamples (75-Item Form of <u>Gumpgookies</u>)	50
Table 4.10--Orthogonally Transformed Factor Loadings Based Upon Factor Analysis of the Item-Test Correlation Coefficients for Ten Subsamples (75-Item Form of <u>Gumpgookies</u>)	52

CHAPTER 5

Table 5.1--Analysis of Variance of <u>Gumpgookies</u> Scores for the Three Highest and Three Lowest Ranked Students on Motivation, for Ten Classes	59
--	----

List of Tables (cont.)

Table 5.2--Analysis of Variance of <u>Gumpgookies</u> Scores for the Three Highest and Three Lowest Ranked Students on Motivation, for Fifteen Classes	60
Table 5.3--Analyses of Variance of <u>Gumpgookies</u> Scores for the Fifteen Per Cent Highest and Fifteen Per Cent Lowest Ranked Students on Motivation	61
Table 5.4--Oblique (Oblimin) Factor Loading Matrix for a Six-Factor Solution for the <u>Children's Self-Concept Index</u> (104 Hawaiian Second-Grade Children, Spring, 1969)	64
Table 5.5--Oblique (Oblimin) Factor Loading Matrix for a Two-Factor Solution for the <u>Children's Self-Concept Index</u> (104 Hawaiian Second-Grade Children, Spring, 1969)	65
Table 5.6--Difficulty Values (Percentages Passing) and Item-Test Point Biserial Correlation Coefficients for the Items in the <u>Children's Self-Concept Index</u> (Hawaiian Second-Graders, N = 104, Spring, 1969) . .	67
Table 5.7--Frequency Distribution of Scores on the <u>Children's Self-Concept Index</u> (Hawaiian Second-Graders, N = 104, Spring, 1969)	68

CHAPTER 6

Table 6.1--Obliquely Rotated Factor Loadings for the 100-Item Group Form of <u>Gumpgookies</u> for Three Grade Groups (First, Second, and Fourth)	78
Table 6.2--Correlations Among the 18 Sets of Oblique Factor Loadings for the Six Factors for Each of Three Grade Groups (First, Second, and Fourth), Based on a 100-Item Group Form of <u>Gumpgookies</u>	84
Table 6.3--Two Sets of Correlated Factor Loadings Across Grade Groups	86
Table 6.4--Orthogonally Rotated Factor Loadings Based Upon Pearson Correlation Coefficients of Obliquely Rotated Factor Loadings for First-, Second-, and Fourth-Graders (100-Item Form of <u>Gumpgookies</u> , N's 212, 223, and 233, Respectively)	88
Table 6.5--Orthogonally Rotated Factor Loadings on Six "Super-Factors" of the Six Initial Orthogonally Rotated Factors for Each of Ten Subsamples (Based on 75-Item <u>Gumpgookies</u>)	96

List of Tables (cont.)

Table 6.6--Initial Factor Numbers for the Positively Loaded Factors Associated With Each Super-Factor for the Ten Subsamples, Together With the Loadings on the Super-Factors	98
Table 6.7--Relation Between Two Initial Super-Factor Solutions, With Associated Numbers of Items	101
Table 6.8--Relation of Orthogonal Six-Factor Solution for Grades 1, 2, and 4 (N = 668) and Orthogonal Six-Factor Solution for Ten Ethnic-Cultural Groups (N = 1,607)	115

CHAPTER 7

Table 7.1--Analysis of Variance of <u>Stanford-Binet</u> Pre- and Post-Test Scores for Three Motivation Classes Combined (N = 59) and Three Comparison Classes Combined (N = 41)	134
Table 7.2--Analysis of Variance of 100-Item <u>Gumpgookies</u> Pre- and Post-Test Scores for Three Motivation Classes Combined (N = 59) and Three Comparison Classes Combined (N = 41)	135
Table 7.3--Analysis of Variance of 75-Item, Randomized Form of <u>Gumpgookies</u> for Three Classes	136
Table 7.4--Analysis of Covariance of Scores for the 75-Item, Randomized Form of <u>Gumpgookies</u> With the Pre-Test 100-Item, Left-Right Form of <u>Gumpgookies</u> as Covariate for Three Classes	137

FOREWORD

The studies reported began in 1965 and have been continually supported by the Education Research and Development Center of the University of Hawaii. Substantial financial resources have also been made available, beginning in 1967, by the Office of Economic Opportunity, for projects entitled "Measurement of Motivation To Achieve in Preschool Children" and "Development of Techniques To Teach Motivation To Achieve to Preschool Children, With Further Refinements of a Measuring Instrument" as well as through the funding of the Head Start Research Center in 1967-68 and 1968-69.

It is hoped that, because of the extended period of time that has been required for this series of studies and the attendant turnover of staff, the authors of the present report will be forgiven for not mentioning all of the many persons who have contributed in one way or another. These include officials of the Honolulu Community Action Program and of the Hawaii Office of Economic Opportunity; directors of Head Start Centers; Head Start teachers and teacher aides; parents of children enrolled in Head Start classes; personnel in various capacities in the Hawaii Department of Education; and directors of and teachers in private preschools. Without the cooperation of these persons, the series of studies reported herein could not have been done.

Some of the persons who have contributed substantially to the design or the numerous analyses reported herein will nevertheless be named:

David G. Ryans, Director of the Education Research and Development Center, lent his ever-ready ideas, encouragement, and administrative support. Other members of the staff of this Center, including Ian Reid and Peter Dunn-Rankin, have also generously contributed technical assistance.

Bonnie L. Ballif, who was formerly on the staff of the Education Research and Development Center and is now at Fordham University, has worked on this series of studies since 1966, as an on-campus employee of the University of Hawaii and as an off-campus employee or in a consultant capacity. She has had a major role in all aspects of the work, including the preparation of this report.

Vidya Bhushan served as statistical consultant with respect to many of the analyses; Paul Horst was an invaluable consultant at a critical stage; Jacqueline Martin and Diane Selser interpreted the design of the curriculum to teachers as well as assisting in testing, coding data, and other work.

Coordinators of testing in locales outside Hawaii included John Gilheney in New York and Utah, Carolyn Stern in Los Angeles, Thelma Harms in San Francisco, Ruth Waugh in Oregon, and Renato Espinosa in Texas. Hannah Herman, the Center's Evaluation Coordinator during the period 1966-69, was often consulted and contributed substantially.

Renato Espinosa, who has recently joined the Education Research and Development Center and serves as Assistant Director of the Head Start Research Center, has guided many of the later analyses, drafted sections of this report, and contributed in many other ways. Nancy Johngren has been responsible for the details of many of the later statistical analyses. Other current or former members of the Center's staff who have contributed to the project in one capacity or another include Patricia Nash, Carole Hodges, Annette Okimoto, Virginia Lerner, Betty Elrod, June Kimura, Phyllis Loveless, Gayle Geiger, Ann Pavelko, Carole Smith, Ann Nicol, Annie Worth, Elizabeth Crooker, Elizabeth Parker, Patricia Hart, Carol Blair, Nina Lauritsen, Martha Phillips,

Grace Diamond, Katherine Payton, Noelani Rickard, Linda Thopy, Mary Choy, Blanche Shigeoka, Linda Nomura, Georgia Sakai, and Fay Agena.

The report has been edited by Lynne Meyer and typed by Yaeko Santoki and Lynette Kapahua.

ABSTRACT

This research on motivation has been focused on (a) development of a measure of motivation of young children to achieve in school, a measure consonant with a five-faceted theory of constituents of motivation; and (b) a curriculum for teaching motivation to preschool children, also grounded in the theory.

The resulting measure, Gumpgookies, is an objective-projective device consisting of dichotomous items. Different formats are used for individual administration to preschool children and for group administration to elementary school children through age nine or 10. The test has been given most recently to 1,607 preschool children in 10 ethnic-cultural groups scattered over the United States and to a total of 668 children in the first, second, and fourth grades in Hawaii.

Item difficulty indices and item-test correlation coefficients have been studied, and data for each group and for various combinations of groups have been factor-analyzed. Reliability estimates have been based principally on KR-20 values, ranging between .80 and .90.

Content validity, as determined by judgment, is in a sense inherent in the construction of items to accord with a theoretical framework. Evidence is presented for factorial validity; and correlations with other variables and with teacher rankings of motivation provide confirmatory evidence of empirical validity.

Problems related to extraneous influences on the factors have been faced head-on: position of the answer for each item; order of presentation of the alternatives for each item; and, to a lesser extent, position of the item in the test. Tentative explanations of these dispositional

sets in relation to substantive interpretations of factors are offered, along with suggestions toward elimination of their influence on composition of factors.

A curriculum for preschool children also had its origins in the theoretical framework that posits five components of motivation. Results of the first attempts to teach motivation to preschool children are regarded as sufficiently promising to warrant intensive efforts. Plans for extending this work to children both younger and older than preschoolers are being pursued. Follow-up of the groups involved in the study is also projected.

CHAPTER 1

PROBLEMS INVESTIGATED

The interdependence of a child's ability to learn and his desire to learn in determining his success in school is undeniable. The relationship between these intellectual and motivational variables, however, is not well understood, nor is the nature of either of these variables singly. Nevertheless, some theoretical and empirical endeavors are beginning and suggest that motivation may be the more crucial of the two, inasmuch as it not only impels action for learning (Sears, 1966) but also determines the kind and amount of learning that takes place and is ultimately responsible for performance, from which all learning is inferred (Hilgard & Bower, 1968).

Explanations of the development of motivation for learning are primarily theoretical. Although the potential to become so motivated possibly can be regarded as innate, such motivation appears to be at least partially determined by psychological and social experiences and environments (Brim, 1960; Henderson, 1967; Smith, 1965). If a causal relationship between specific conditions and motivation to achieve can be demonstrated, these conditions can then be used to develop such motivation wherever it may be deficient.

There are children who display a progressive lack of achievement-oriented behavior. They have been described as limited in curiosity and interest, expressing self-devaluating attitudes toward their own achievement; lacking pleasure or interest in accomplishment; having no discernible drive towards goals or their attainment; and exhibiting low initiative, autonomy, and independence in self-activated play (Mattick, 1965).

It appears that the first step in the education of any child who meets this description should be to enhance his desire to learn in school (Kagan, 1966). It has been said that increasing his motivation is not only important (Savitsky, 1965); it is paramount, all-pervasive, and critical (Gordon & Wilkerson, 1966). And furthermore, it has been contended that in such cases, efforts to improve motivation for learning should constitute a major part of the lesson time and emphasis (Savitsky, 1965). Accomplishing this task is inescapably a complex problem. Admittedly, the school constitutes only a fraction of the child's total environment. Nevertheless, there may be effective procedures that can be carried out in the classroom if planned intervention occurs early in life when fundamental patterns of responding are being learned (Torrance, 1965).

In order to identify instructional techniques effective in increasing motivation for learning, however, some means of assessing this motivation is required. Each new curricular approach sorely needs painstaking investigation of what does work and what does not (Hunt, 1967); creative endeavors must be checked by precise measurement of their outcomes (McClelland, 1958). Prerequisite to research for discovering effective instructional techniques, therefore, is a valid measure of motivation to achieve. The lack of suitable instruments of this nature is obvious and has severely limited significant findings in this area (Katz, 1967; McClelland, 1958). Hence, it seemed imperative that the resources and energies available for these studies be focused initially on the measurement problem, since real classroom advances await methodological development (McClelland, 1958).

With this in mind, the general objectives of this research concentrated first on the further development of a new measure of motivation to achieve, Gumpgookies, and secondly, on the design of new instructional techniques to teach motivation to achieve to preschool children. A brief description of more specific objectives and the procedures used in their attainment follows.

Some preliminary work on creating a test for measuring motivation to achieve in preschool children had already been completed (Adkins & Ballif, 1967), and an initial trial of this instrument (Gumpgookies) looked promising (Adkins & Ballif, 1968). Inasmuch as all of the earlier research on Gumpgookies had been carried out on children in Hawaii, it was felt that its effectiveness needed to be investigated for children of preschool age from a variety of ethnic and cultural environments. This, then, became the first objective of the present series of studies.

Arrangements were made with professional personnel in selected locations to direct the administration of Gumpgookies to specified groups of children. When it was decided by the National Head Start Evaluation and Research Centers to use Gumpgookies in their cross-country evaluation of Head Start in 1968-69, plans were altered somewhat. Since data from these centers eventually would be available and would contain scores for children from some of the populations originally identified for investigation, this research extended its design to populations beyond those enumerated in the original proposal to include samples of preschool children from three specific religious backgrounds-- Jewish, Mormon, and Catholic. The field-testing of Gumpgookies was temporarily postponed, however, when an unanticipated problem was encountered: responses of some children were clearly influenced by

the spatial position of the gumpgookie figures and possibly by the order in which they were described. It was then necessary to examine these position sets in greater depth and to revise the format in an effort to alleviate this tendency. Accordingly, a revised form of Gumpgookies, limited to 75 items, reduced in verbal difficulty, and with radically new spatial positions of figures was completed and administered as planned.

Appropriate statistical techniques were selected and adapted to computer programs; and test statistics, item statistics, and factor analyses were assembled for each subsample. These data will be used in subsequent revisions of the test. In addition, the Head Start Evaluation and Research Centers were asked to rate the usefulness of each item in the 100-item form and to comment on its appropriateness for specific groups of Head Start children. These ratings and comments not only assisted in the reduction of the verbal difficulty and the length of the test, but were of distinct value in identifying and eliminating aspects of the items subject to cultural differences in interpretation.

It was also desired to explore the potential usefulness of Gumpgookies beyond the preschool by giving it an initial trial on school-age children. Two separate group-test formats were developed, pilot-tested, and tried out on samples of approximately 200 children in each of the first, second, and fourth grades. The children in each of these samples represented high, middle, and low socioeconomic areas. Item analysis and factor analysis procedures were applied to these data, and attempts were made to determine the similarities and differences of the factors between grades.

The inter-relations among Gumpgookies scores, cognitive aptitude scores, and objective tests of academic achievement for the school-age children were determined. The following test data were available in

the cumulative folders: the Sequential Test of Educational Progress and the School and College Ability Test for the fourth-grade classes; and the California Test of Mental Maturity and the California Reading Test (Primary Form W) for the second-grade classes. The relationship of Gumpgookies to the Children's Self-Concept Index was also investigated.

Development of new instructional techniques to teach motivation to preschool children constituted a major goal. A variety of classroom procedures were designed that would capitalize on the ongoing human interaction between teacher and children, and that would provide specific daily activities for teaching the response constituents of motivation to achieve. This program emphasized the training of teachers in presenting the appropriate stimulus conditions in order to help children make the desired overt responses.

The relative effectiveness of these instructional techniques in three experimental classrooms was studied by comparing their gains with those of three control classrooms. These comparisons were made on the basis of pre- and post-test data, along with periodic evaluations of the teachers' presentation of the curriculum.

Projections for this series of studies also entailed preparation for follow-up research on the effects of the various instructional techniques on academic success in kindergarten and first grade. All children participating in this research were identified and their scores on a number of appropriate measures, including Gumpgookies and the Stanford-Binet, recorded. Background information on each child was also collected and compiled, making it available for follow-up investigations.

CHAPTER 2

THEORETICAL ORIENTATION

Investigation of motivation to achieve in learning required initial clarification of the construct itself. A logical source for this needed insight was relevant theoretical and empirical literature focused on achievement motivation. Much of this literature, understandably, is still speculative; and motivation of any sort can only be inferred from behavior (Cofer & Appley, 1964). Hence a distillation of concrete ideas from inferences based upon such an indirect process has not been an easy task.

Nevertheless, running throughout the various positions and investigations are several recurring themes that have influenced the theoretical framework within which the present series of studies have been formulated. To assign full credit to the numerous sources that have helped to create and refine the ideas to be discussed is scarcely possible; however, indebtedness to numerous pioneering efforts will be acknowledged, through citations of references, as appropriate, in the presentation of a conceptualization of motivation. At this stage in its development, no attempt will be made to state a completely thorough or in any sense final theoretical position or set of hypotheses. Many aspects of the relationships among the different elements and their developmental aspects are as yet unknown.

The hypothetical components, however, have seemed appropriate to their purpose, i.e., to serve as a basis for conceiving of motivational variables and hence for giving direction to continuing research. If the lines of investigation pursued eventuate in empirical substantiation of

the fruitfulness of the conception of motivation to be presented--as will later be suggested is indeed the case--then there will be an impetus for a more complete formulation. It is recognized that the ultimate development of integrating principles must await the establishment of a firm empirical base.

At the present stage in this series of studies, motivation to achieve is being conceptualized as a hypothetical construct that explains aspects of achievement-oriented behavior that are not attributable to intellectual abilities, whether innate or learned. Motivation to achieve appears to be determined by a combination of attitudes, feelings, or expectations. These may be conceptualized as unobservable phenomena not unlike observable phenomena in that both are subject to lawful determination under appropriate causal conditions. Such unobservable phenomena may appropriately be regarded as covert responses.

Doob (1947) considered attitudes to be psychological responses, evoked by a variety of stimulus patterns as a result of previous learning or of gradients of generalization and discrimination, and themselves cue- and drive-producing.

Certain types of these learned covert responses appear to accompany instances of achievement-oriented behavior from which motivation to achieve is inferred. Attempts to categorize these responses first divided them into seven, then into six, and finally into five major areas. Each of these five types of covert responses is hypothesized to be an essential component of motivation to achieve. Each of these responses may belong to a family of responses of which one or another could be aroused by a similar set of cues (Hardy, 1964). Neither their exact

nature nor their dynamic interaction is markedly clear. It does seem, however, that all five types of responses are essential if motivation to achieve is to occur.

Because the current studies have been focused on motivation as related to school achievement, especially that of young children, the discussion of the components of motivation is often presented in terms reflecting this concern with learning of young children in the school environment. It should be recognized, however, that the components that have been identified for special consideration are general, in the sense that they are applicable to motivation to achieve in any context.

The first of these covert responses is rather widely designated as expecting affective or hedonic change (McClelland, 1951; Young, 1959; Hardy, 1964). Some concepts of affect and affective arousal are fundamental in the theories of Atkinson (1954), McClelland (1953), Olds (1955), and Young (1955). Atkinson (1965) discusses this response in terms of the relative attractiveness of success or failure for a particular activity and refers to it as the incentive value of success or failure. Hardy (1964) states that motives are based upon learned expectation of affective change. He suggests that there is a positive relationship between the amplitude of the affective change and the intensity of the resulting motive, that the more immediate the affective change the more attractive it will be, and that the more permanent the affective change the more compelling the motivation.

The source of affect has been considered in Hebb's (1949) neurological model, in Helson's (1947) notion of adaptation level, and in McClelland's (1953) discrepancy hypothesis. (See also Haber, 1958.) The research related to these and other investigations, while fascinating, is not so

pertinent to the immediate goals of this theoretical orientation as is the idea that all behavior is motivated by expectations of affective change or increase in positive affect.

The young child must expect that if he engages in achievement-oriented activity within the school setting his life will be made more pleasant--otherwise, he will not be motivated to engage in such activity. Although the nature of the more pleasant life may or may not be within the school context, it must be an end for which achievement-oriented behavior is seen to be the means. When this means-to-end relationship exists, motivation to achieve may appear. Thus a necessary condition is that the child have a positive orientation toward school and achievement-oriented activities; he must expect enjoyment from working and accomplishing within the school setting.

The second hypothesized covert response constituent of motivation to achieve is conceptual in nature, focusing specifically on the concept of self as an achiever in learning. This response is seen by Brookover, Erickson, and Joiner (1969) as a process of characterizing oneself within the academic role as a result of learned interactions with others. Combs and Soper (1963) also believe that behavior is a function of how the individual is perceiving himself within the world in which he is operating. They conclude that perceptions about the self are crucial in the causation of behavior and that the feeling of personal adequacy is of pervasive importance in the child's perceptual organization and functioning in the classroom.

Hence motivation appears to be accompanied by the individual's expecting to be successful in the particular situation. Atkinson (1965) proposes that the resultant motivation is in part a function of the

strength of the individual's expectancy that his performance of a task will be followed by success or of the individual's subjective probability of attaining his goals.

If a young child sees himself as being successful in achieving within the learning environment, he appears to be motivated for achievement and engages in appropriate achievement-oriented activities, approaching them with confidence in his ability to excel.

Expecting positive affect from achieving in school and having confidence in being able to achieve are important but not sufficient conditions to result in motivation to achieve. A third component that is regarded as essential arises from the direction or purposiveness of behavior implied in the concept of motivation itself. It is, in essence, the setting up of purposes for the self-direction of behavior. These goals often go beyond the immediate moment and suggest implications for other times, places, and people (Todd, Terrell, & Frank, 1962). Demonstrably, however, specific performance goals also increase motivation (Bryan & Locke, 1967). This responsiveness to and setting up of goals is present whenever motivation to achieve is apparent.

Closely related to the purposiveness of behavior is the knowledge of instrumental steps that will be effective in accomplishing the purposes. A child could expect positive affect from achieving in school, see himself as an achiever, and be able to set up purposes to direct his behavior towards achievement but still not ever exhibit achievement-motivated behavior unless he knows what to do to accomplish his goals. The first instrumental step toward any purpose is the realization of personal responsibility for action and of personal control over outcomes. An individual must believe that some action on his part helps or is

required to result in the desired goal. This idea is present in Atkinson's (1965) model of risk-taking behavior in that it does not apply to matters of chance, i.e., lack of individual control over outcomes. Instrumental behavior is a necessary part of a motivational expectation (Hardy, 1964). Recognizing the relationship between means and an end and knowing the appropriate means to the end are also characteristic of individuals motivated to achieve (Raynor, 1967; Todd, Terrell, & Frank, 1962). In addition to the child's belief that he is the causal determinant of the outcome of his behavior, therefore, he must also know that he should autonomously initiate work activity that will be instrumental in accomplishing his purposes. These are all cognitive realizations or responses.

The last of the five covert responses hypothesized to be components of motivation to achieve is self-evaluation. A positive self-concept itself is not sufficient to cause achievement-motivated behavior. Some process of self-assessment or self-evaluation is also essential (Brookover, Erickson, & Joiner, 1969). Katz (1967) states that a child's capacity for sustained academic effort depends heavily upon an internalized mechanism of affect-mediated self-evaluation. Several investigations conclude that autonomous evaluations correlate positively with independence and achievement motivation (Argyle & Robinson, 1962; Dreyer, 1966; Martine, 1956; Pedersen, 1965).

This process of self-evaluation requires not only the presence of some internal standard of excellence but also the ability to compare actual performance with this standard and to discriminate between the two. In this way, further direction is provided for future behavior. This response is being referred to as an ethical response. Note that the term ethical, as used here and throughout this report, is intended

to refer to self-evaluation of the relative goodness or badness of one's performance that is directed toward certain goals, not to value judgments of the intrinsic goodness or badness of the goals themselves.

Achievement-motivated behavior is a result of the dynamic interaction of these responses. Motivation to achieve in school will be evident only when a child expects that achieving in school will be pleasant; when he thinks that he can achieve in school; when he can set up his own purposes to achieve; when he knows the instrumental steps that will lead to his achievement; and, finally, when he can evaluate his own performance against some internalized standard of excellence.

The relative power of these responses in determining actual achievement and their dynamic interaction need to be investigated. Their respective strengths may vary inter- and intra-personally. It is hypothesized that all must be present in some degree, however, in order for motivation to achieve to be evidenced in behavior.

Inherent in the conceptualization of these components as covert responses is the idea that they are learned. As previously pointed out, Doob (1947) considered attitudes to be learned covert responses, and Malzman (1955) considered thinking equivalent to a new combination of habit strengths produced by mediated generalizations. He further suggested that some of the elementary laws derived from conditioning will lead to the development of composition laws operating in human problem-solving. It has also been suggested that the actual process of learning covert responses such as problem-solving and attitudes must be described. Furthermore, it should be experimentally demonstrated that formation and change in these responses can be adequately explained by the principles established for learning of other responses (Staats, Staats, & Heard, 1960).

If these covert responses are indeed learned, they should be teachable in the sense that they may be increased in predictable directions by appropriate control of the independent variables. Apparently some school situations already are effective to some degree in increasing these constituents of motivation (Zigler & Butterfield, 1968).

A great deal of research is needed to find out what principles are involved in the learning of each type of response. The problems are inescapably complex. The interactions of any one individual with a multitude of people, values, and experiences are at first glance of an imponderable magnitude. Fortunately, however, each individual selects from this totality certain of the presented stimuli to which to attend at any one time. The causes of behavior, therefore, are imbedded in the perceptual organization of the behavior, the individual's perceptual field (Combs & Soper, 1963).

It is the position of this research that those covert responses that constitute motivation to achieve within the child's perceptual field can increase in strength. Motivation can be changed (McClelland, 1965). Atkinson (1965) suggests that, if the individual succeeds, his expectancy of success increases and that this expectancy of success is a motivational variable that can be manipulated experimentally. He further argues that manipulation of his expectancy is the most feasible means of bringing about changes in achievement-oriented motivation in schools. Moreover, Hardy (1964) states that motives are a function of a variety of cues associated with affective change, together with the associative strength of those cues to the affect. Perhaps the essence of these comments is that, through contiguous associations of experiences and affective change, a child can learn the response components of motivation to achieve.

These associations can be established only if the stimuli involved become part of the child's perceptual field. One promising means to enter this field is through imitative learning. Bandura (1962) claims that this method is effective and still places primary emphasis on contiguous sensory stimulation as a sufficient condition for the acquisition of most forms of matching responses.

Flanders (1968) has extended Bandura's concept of imitation to include verbalization that may affect the acceptance of the model by increasing knowledge. Empirical evidence does support the idea that both direct training and observation of models are more effective when the training agent verbalizes the contingency for his performance (Liebert & Allen, 1967) and that children who can verbalize their awareness of certain covert responses exhibit greater independent ambition (Gilbert, 1969). The constant encouragement of verbalization, therefore, may be an additional means of entering the child's perceptual field and hence of more effectively teaching the responses in question.

Hardy (1964) predicts that affective expectations increase each time they are confirmed.

Coopersmith (1969) outlines three conditions that he feels are associated with the development of high self-esteem. The first, acceptance, is defined by an expression of warmth, interest, and concern for the individual's well-being by persons significant to him. The second condition is that of clearly defined limits and goals, accompanied by relatively high demands and expectations for performance. The third and last condition is respectful treatment and latitude for individuality for persons who abide by established limits.

Brookover and associates (Brookover, Erickson, & Joiner, 1969) concur with Coopersmith and emphasize that one's concept of his own ability changes over time in relation to shifts in perceived evaluations made by significant others. They empirically showed that change in the evaluations made by parents significantly affected self-concepts related to ability and grade-point average in ninth-grade students. Similar changes in the evaluations of counselors and learning experts did not have similar effects.

That purposes can be learned is also possible (Rasey & Menge, 1956). Through verbalization and observation of appropriate models, instrumental steps in obtaining goals may also be taught.

A number of studies are providing evidence for establishment of internal standards of excellence in performance and hence self-reinforcement procedures through the use of appropriate models (Bandura, Grusec, & Menlove, 1967; Bandura & Kupers, 1964; Bandura & Whalen, 1966; Herbert, Gelfand, & Hartmann, 1969; Liebert & Ora, 1968; Mischel & Liebert, 1966; Ofstad, 1967).

The setting for this research, therefore, has been a hopefully sound, but certainly still tentative and very possibly incomplete, structure. Motivation to achieve is viewed as a dynamic combination of learned responses: expectation of affect, conceptualization of self, formulation of purposes, knowledge of instrumental steps to accomplish purposes, and evaluation of one's own performance. It is suggested that these responses are learned and therefore may be taught within controlled environments.

CHAPTER 3

EARLY METHODOLOGICAL CONSIDERATIONS IN THE DEVELOPMENT OF GUMPGOOKIES

To develop a testing procedure that would not only accurately measure the evasive components of motivation to achieve but also be effective within the limited response repertoire of preschool children was admittedly a major undertaking. Probably the most influential approach to measuring achievement motivation has been the work of McClelland and his associates (McClelland, Atkinson & Clark, 1953), who used fantasy as the medium through which themes, needs, and goals are scored for achievement content. (See also Cofar and Appley, 1964.) Despite the appeal of this idea, complications arise when such a procedure is used with very young children. Many preschoolers tend to withdraw in the testing situation itself. The majority lack verbal skills needed to describe their fantasies. Moreover, the absence of universal child-rearing practices may mean that young children have not been exposed to any uniform series of experiences (Anastasi, 1954). As a result, both their understanding of picture stimuli and the content of their fantasy may be limited.

Accordingly, preliminary searching for a method of measurement that would meet both of these criteria was extensive and included a great variety of techniques and formats. Efforts began by asking teachers to rate the behavior of their children along nine-point scales for each of 10 personality variables believed to be relevant to motivation to achieve (Digman, 1965). These scales were tried first without and then with written descriptions of specific behavior to clarify each interval along the scale. As has characteristically been the

case, the teacher ratings were found to be time-consuming as well as subject to rater idiosyncracies and to criterion contamination. They provided opportunity for only very crude estimates of gains reflecting progress.

Structured situations were then designed to quantify the child's reactions to opportunities to achieve in his everyday activities. These behavioral tests, too, had their shortcomings. They were difficult and expensive to construct, plagued by the situation-specific reaction tendencies of four-year-olds, and appropriate for measuring only a limited number of the hypothesized response constituents of motivation to achieve.

A further endeavor investigated role-playing in achievement and in self-evaluation while achieving. Children were asked to make believe that they were in school and instructed to do their best on as many of the school-type tasks as they wanted to do. Although it appeared that the role-playing situation may have influenced initial responding, the variability in the attraction of specific tasks seemed to be a predominant factor in determining later response patterns.

Pictorial semantic differentials with illustrated bipolar adjectives were tried out as a possible means to scale the affective and evaluative meaning of school- and achievement-oriented objects and activities. The children, however, appeared to have some difficulty in relating the words to the pictures, responding impulsively to the illustrations themselves and not to their symbolic representation.

In a second attempt to quantify the affective responses contributing to motivation to achieve, a procedure for measuring feelings and attitudes towards achievement-oriented activities and objects was constructed by

drawing a face on each of three boxes. These faces, depicting varying degrees of happiness and sadness, served as visual scales of expected affect. A number of different games were created that required the child to express his feelings by identifying the appropriate face on this scale with an illustrated activity or object. Unfortunately, however, young children do not appear to be able to discriminate abstractly and categorize into three separate degrees.

Efforts to simplify this task resulted in trying out a variety of paired-comparison techniques. The most promising was one in which the child was shown a pair of pictures and asked to point to the activity he preferred doing. All the children seemed to understand and perform accordingly. Subsequent questioning revealed that they were also apparently making valid comparative judgments in the forced-choice decisions. The procedure itself became monotonous, however, and the children lost interest in it and withdrew.

With hope that it would be possible to increase interest in a paired-comparison procedure, a story was written describing the daily life of a four-year-old child and including a limited number of illustrated items with dichotomous alternatives. This approach succeeded in sustaining interest and attention, but required too many descriptive words and provided too few opportunities for choices per unit of time invested. In addition, identification with human character was at times threatening to some children.

The results of these studies and complete discussions of their findings are available in a previous report (Adkins & Ballif, 1967). From these initial endeavors, however, sufficient direction was obtained for the creation of a new measure of motivation to achieve, Gumpgookies.

Gumpgookies: 200-Item, Left-Right Form

Gumpgookies is an objective-projective technique that requires a choice between two types of alternative behavior portrayed in pictures and accompanying verbal descriptions. It appears in a story format and centers around the activities of imaginary little figures called gumpgookies. The gumpgookies behave in ways intended to show differences in motivation to achieve in activities appropriate for preschool children and possibly for older children up to about age 10. Each item consists of two little gumpgookies in a semi-structured situation. The child is told that he has his own gumpgookie and that, although it looks like all others, it follows the child and behaves as he behaves--it likes what he likes, and it does what he does. As the examiner reads the story and points to each gumpgookie in turn as it is described, the child is asked to find and point to his own gumpgookie. Two illustrative items are as follows:

Here are two gumpgookies getting up in the morning.

This one thinks it will be a good day.

This one thinks it will be a bad day.

Which is your gumpgookie?

This gumpgookie does what it wants to.

This gumpgookie does things well.

Which is your gumpgookie?

From a pool of some 300 items, 200 items were selected for the preliminary form. It should be noted that in this first form, the two gumpgookies appeared side by side, with the left one being described first and then the right one for each item. This instrument was administered in two sittings to 182 children--114 from Head Start.

classrooms and 68 from private preschools. The age range was from 49 to 80 months, and the average age was 62 months. Approximately 90 of the Head Start children were selected by pooling judgments of the teachers and their two aides as to the child most motivated to achieve and the child least motivated to achieve for each of 55 Head Start classrooms.

For these data, measures of the relation of each item with the score on the total 200 items, as well as discrimination indices for the external criterion (i.e., high versus low motivation), were determined.

In addition, the matrix of inter-item phi correlation coefficients was factored by the principal-axes method and the factor matrix rotated to an oblique simple structure, using a biquartimin solution with gamma equal to .5.

The computer program, however, did not specify number of factors to be extracted, and it was found that the eigenvalues* had not decreased to unity even when as many as 20 factors had been extracted. Little hope was justified that as many as 20 factors could be interpreted; hence the number of factors to be removed was arbitrarily set at six or seven and, for some solutions, at three. This failure of the eigenvalues to approach unity after what seemed to be a reasonable number of factors had been removed highlighted certain problems connected with factor analyses of binary data.

The rotation of this initial set of factors permitted only extremely tentative identification of seven factors, which in turn provided only

*It has been suggested that the number of common factors in a matrix is equal to the number of eigenvalues greater than one. (Kaiser, 1960).

limited evidence for the hypothesized response constituents of motivation to achieve. The inter-correlations among the seven factors were also factored, yielding a three-factor, second-order matrix that was also rotated to an oblique simple structure. The second-order factor analysis provided what seemed to be a somewhat clearer three-factor structure. In view of subsequent developments, however, interpretation of these factors will not be presented here but has been made and is available (Adkins & Ballif, 1968).

Gumpgookies: 100-Item, Left-Right Form

At this point, Gumpgookies was revised to consist of 100 items and was administered in one sitting to a new sample of 330 children. The data were again analyzed in terms of basic test statistics; and although factor-analysis techniques were applied, a number of alternative approaches were also pursued. One was a type of cluster analysis designed to yield clusters with maximum KR-20 reliability estimates.*

The factors and clusters were not expected to coincide perfectly; nevertheless, the first cluster contained variables primarily loaded on three of the factors, and the second cluster contained variables coinciding with those loaded heavily in two other factors. In general, the results of this cluster analysis were rather similar to those of ordinary factor-analytic methods. Moreover, anomalous results, such as negative reliability coefficients, sometimes occurred, and the question of a modification of the program that was possibly needed had not been resolved. Hence this technique eventually was abandoned.

*A computer program was made available by Joseph Klock.

Another alternative approach, dimensional analysis of binary data, was brought to the attention of the investigators by Ledyard Tucker (Congor, 1968). Careful consideration of this method, however, led to the conclusion that it could be expected to lead to about the same results as the more traditional factor-analytic techniques. Hence it was not pursued.

The prospect of difficulty factors in the analysis of binary data was not unknown to the investigators. A number of references on such factors were consulted, beginning with Sir Cyril Burt's early mention of them and including later papers (Burt, 1941; Carroll, 1945; Dingman, 1958; Ferguson, 1941; Gourlay, 1951; Guilford, 1941). After consideration of the problem and discussions with Paul Horst, the decision with respect to it was to proceed with factor-analytic techniques and attempt to interpret factors that would transcend difficulty factors, even though beclouded by them to an extent.

Let us return to the early analyses of the 200-item, individually administered form of the test. Primarily on the basis of difficulty values and item-test correlations, but also partly on the basis of factor analyses, the length of the test was reduced from 200 items to 100.

Although the answer key for the original 200-item form had been determined in a random order, it was apparent that the key for the resulting 100-item form warranted revision because an unusually large number of the items selected had answers corresponding to illustrations on the right-hand side of the page. These also coincided with the verbal descriptions read last by the examiner. Although this discovery was startling, it was not inconsistent with the fact that events that are relatively improbable do indeed occur, with a predictable relative

frequency. Although suspicions had been aroused, vacation periods and demands for a revised form of the test were imminent. Accordingly, the key for the 100 items selected for the revised form was randomized.

Further study of factor and cluster analyses of the data on the 200-item form and of the first 100-item form soon revealed curious problems. Certain factors or clusters had most keyed answers in the right-hand position, others in the left-hand position. Note that those keyed answers at the left had been read first, those at the right last. Hence, with the test format used, the left-right and primacy-recency influences were inextricably confounded.

Inspection of answer sheets for individual subjects did not reveal this problem. The children whose results were analyzed did not display predominant and consistent tendencies to choose one answer position over the other. Apparently what was happening was that when the cognitive difficulty of the items or the required attention span became too great for a particular young child, he temporarily responded according to a position set or a primacy-recency set. However, when he was presented with an item well within his comprehension or when his attention had been directed back to the test tasks, his temporary set was interrupted. For later items, he might resort to responses determined by his underlying sets.

Three principal approaches were pursued in efforts to understand this problem. One was to divide the answer sheets of the subjects into two groups--one that did and one that did not differ significantly from the number of runs (successive responses of right or left answers) appearing in the answer key. Data for the two groups of subjects were then separately factor-analyzed. Without presentation of the agonizing

details of the analyses, it must be reported that the outcomes were inconclusive. With a number of factors and hence a relatively small number of items clearly associated with each, presence of position or primacy-recency factors for subjects exceeding the runs criterion and absence of such factors for subjects within the expectation of runs for the key were not definitely demonstrable. Although much labor had gone into this enterprise, results were rather ambiguous. The most probable explanation of them was that the statistical criterion used to separate the subjects into those susceptible and those not susceptible to runs were not well adapted to detections of the subtle psychological influences that determine what on the surface appear to be erratic shifts of set among preschoolers, given the original format and nature of this particular test, coupled with other problems associated with binary data.

A second attack on the problem of set factors yielded more definitive results. After much trial and error, artificial score matrices, with randomly assigned, equal numbers of answers in each position, were constructed for 24, 30, or 36 subjects.

The answer patterns and item inter-relationships were designed so as to yield two factors, in some cases very clear ones and in other cases weaker ones. Then the original answer patterns were overlaid with complete position preferences (or the equivalent primacy or recency preferences) for varying numbers of subjects. In the case of the strongly determined factors, the imposition of position preferences for roughly a fourth of the subjects altered the original factor structure to "position" factors only; i.e., the resulting factors had answers appearing in only a single position. With weaker initial factor structure, overlaying

position preferences on the answers of even a smaller fraction of the subjects (perhaps a fifth or a sixth) shifted the factors to dominance by answer position.

Even though the straightforward nature of the shifts in the answer patterns in the matrices analyzed by the foregoing means differs from the less easily discernible patterns characteristic of the responses of the four-year-old children on whom the original work had been done, this second approach confirmed that the phenomenon of position factors had to be contended with and, if not overcome, at least somehow be taken into account.

A third method that confirmed the operation of position preferences for the four- and five-year-old subjects was rearrangement of the inter-item matrix of phi coefficients so that the 51 items with correct answers at the right appeared together and the 49 with answers at the left appeared together. Almost without exception, the mean coefficients of items with other items having the same answer position were positive; with other items having the reverse answer position, negative. The mean positive coefficients were almost uniformly larger than the mean negative ones, however.

The fact that there were more than one right and more than one left factors indicated that some content variables were involved. This optimism was bolstered by the fact that many of the items did discriminate between children selected by teachers as having high and low motivation. Moreover, mean total scores of four-year-old Head Start children in a group composed of those selected by the teachers as the three most highly motivated and those selected as the three least highly motivated differed significantly. Further confirmation lies in the fact

that score distributions, even for the youngest groups, did not fall equally below and above a score equivalent to 50% of the items but started at or near the 50% score and progressed upwards. And, in general, mean scores increased with increasing age. The first calculation of the correlation of Gumpgookies scores with IQ yielded a significant r of .31, ($N = 131$, $p < .01$), which again was interpreted to mean that factors other than chance were operating.

Before this problem had been faced head-on, the 100-item form that resulted from the culling of 100 items from the first 200 items was in demand for use as an instrument in the 1968-69 national evaluation of Head Start. The plan was to use it as both a pre-test and a post-test. The reactions of many examiners from various ones of the then 13 Head Start Evaluation and Research Centers were to the effect that the test was too long and that the cognitive difficulty of many items was too great for a number of the younger children from economically deprived homes. With a view to taking into account or at least reacting to these and other, more idiosyncratic objections to particular items or groups of items, the length of the pre-test was cut to 55 items for the post-test for the national evaluation sample. By this time it was clear that certain items could have been worded more felicitously and that the illustrations for some could have been improved in order to eliminate additional extraneous influences. Nevertheless, exigencies of time precluded any revision for the nationwide post-testing of sample Head Start classes except to reduce the test length to 55 items. The plan, then, was that pre-test scores on these same 55 items could be compared with post-test scores; moreover, the comparison could be based upon norms for subjects spaced by one- or two-month age intervals, as

determined from data for the pre-tests rescored for just the 55 items common to the post-test. The method of determining the transformed scores and the resulting table of norms are presented in the Appendix. A wealth of data resulting from the national evaluation study remains to be studied.

Gumpgookies: Randomized Forms

In the meantime, for the projected testing of several ethnic-cultural groups that had been scheduled for the spring of 1969, it was possible to revise the 100-item test further in several ways: (a) the positions of the illustrations were no longer to be confined largely to left and right but could include up and down, lower left and upper right, and upper left and lower right; (b) the order of description of the figures could be randomly determined; (c) the answer positions could again be completely randomized, taking into account both position of the illustration for the keyed answer and the order of presentation; (d) the wording of many items could be simplified to reduce cognitive or verbal difficulty; (e) certain items that were objectionable for one reason or another could be removed; and (f) the test could be shortened, although it was still desired to have enough items left for full exploration of various ideas about motivation in young children. Detailed comments on and ratings of each item made available by examiners from the various Evaluation and Research Centers were studied in detail, as well as results of all data on item difficulty, item-test correlation, and factor and cluster analyses.

The 75-item, completely randomized form that resulted from the several considerations above was used for the spring, 1969, testing of 10 different ethnic-cultural groups. An intermediate 100-item form,

also completely randomized, had been previously adapted for group testing of older children. One such form, with five color-coded items per page, is usable for non-readers (first- and second-graders, and possibly kindergarten children). A second such form, not requiring the examiner to read each item and cue it to a color, is usable for fourth-graders (and perhaps third-graders). These are the two forms used for the testing of first-, second-, and fourth-graders and reported on hereafter.

CHAPTER 4

BASIC TEST AND ITEM STATISTICS FOR GUMPGOOKIES

This chapter is divided into two sections, one dealing with test data for the 100-item randomized group form of Gumpgookies that was administered to Hawaiian children in first, second, and fourth grades, the other with data collected on younger children for 10 ethnic-cultural groups.

Data for the 100-item Randomized Group Form of Gumpgookies for Three Grades

The 100-item randomized group form of Gumpgookies was administered to 212 first-graders, 223 second-graders, and 233 fourth-graders in three schools in Oahu. These schools represent a cross section of socioeconomic levels. The children in one of the schools came from an area of low income and high density, those in a second school came from families of a low- to a middle-income level, and those in the third school from families that were predominantly of a middle-income level.

For first- and second-graders the items were read by the examiner, while fourth-graders read the text by themselves.

Table 4.1 contains the means, standard deviations, measures of skewness and kurtosis, KR-20 reliability estimates, ranges, and numbers of cases. The distribution of scores for the first-graders is negatively skewed and leptokurtic.

Examination of the item-test correlations and difficulty indices (percentages of subjects choosing the motivated gumpgookie) of the items presented in Table 4.2 indicates that several elicited differential

TABLE 4.1

Distribution Statistics for the
100-Item Group Form of Gumpgookies for Three Grades

	First Grade	Second Grade	Fourth Grade
Mean S.E.	83.03 .74	84.44 .63	82.21 .59
Standard Deviation S.E.	10.80 1.16	9.36 .91	9.06 .56
Skewness S.E.	-2.52* .17	-1.99 .16	-1.03 .16
Kurtosis S.E.	7.81* .33	6.36 .32	1.57 .32
KR-20	.90	.88	.86
Sample Size	212	223	233
Range Low High	26 96	33 96	43 97

* $p < .05$

TABLE 4.2

Difficulty Values (Percentages Passing, D) and Item-Test Point-Biserial Correlation Coefficients (C) for the Items in the 100-Item Group Form of Gumpgookies (First Grade N = 212; Second Grade N = 223; Fourth Grade N = 233)

	First Grade		Second Grade		Fourth Grade	
Item	D	C	D	C	D	C
1	90	36	95	34	94	20
2	76	21	84	18	96	23
3	90	12	97	00	95	-01
4	84	21	92	20	93	15
5	71	39	75	27	81	28
6	88	27	83	48	73	50
7	94	28	92	35	95	22
8	23	-03	29	03	63	05
9	48	17	61	04	85	13
10	90	24	85	21	99	33
11	89	45	96	38	94	25
12	81	10	87	31	97	30
13	93	24	95	33	93	30
14	95	47	97	39	97	15
15	86	47	89	27	90	13
16	90	48	94	47	86	49
17	82	29	81	40	65	47
18	89	52	93	18	94	41
19	84	09	89	25	88	00
20	84	37	77	45	64	42
21	94	66	98	32	98	24
22	75	31	78	33	76	09
23	90	57	91	38	85	43
24	66	22	83	27	54	24
25	75	32	93	35	95	25
26	92	57	95	32	94	28
27	91	39	83	56	54	60
28	74	29	79	11	85	20
29	89	43	85	54	88	42
30	84	25	80	16	73	-05
31	97	63	97	29	98	34
32	91	49	91	39	85	42
33	92	53	90	42	94	26
34	49	13	61	09	58	30
35	85	39	81	41	69	57
36	63	21	73	06	84	41
37	93	52	89	30	73	30
38	50	24	45	24	43	24
39	80	41	85	21	76	09
40	86	48	86	20	88	35

TABLE 4.2 (cont.)

	First Grade		Second Grade		Fourth Grade	
Item	D	C	D	C	D	C
41	58	09	57	27	73	26
42	84	18	98	42	96	16
43	71	27	63	31	61	37
44	96	14	93	29	77	38
45	92	54	95	36	96	30
46	93	57	99	26	100	04
47	95	63	97	18	97	19
48	72	32	68	45	47	51
49	73	06	88	17	94	10
50	81	24	83	17	88	19
51	82	29	71	43	53	49
52	96	36	92	47	86	28
53	82	25	66	41	56	43
54	76	18	76	29	61	12
55	56	01	50	00	22	22
56	88	43	83	17	95	13
57	85	47	82	34	67	45
58	90	24	90	18	89	01
59	94	50	99	38	98	07
60	92	60	92	44	88	34
61	64	36	61	19	62	09
62	86	35	92	44	91	32
63	90	56	81	51	65	54
64	33	-07	53	01	91	14
65	58	-02	65	27	68	29
66	90	40	92	44	96	29
67	82	14	91	09	84	26
68	93	67	93	36	94	42
69	93	48	95	43	91	36
70	84	36	83	39	82	34
71	91	56	92	45	93	15
72	81	22	91	17	95	15
73	93	39	96	28	96	11
74	90	54	90	30	96	30
75	90	35	93	50	91	50
76	80	27	55	-00	35	06
77	95	21	99	18	98	17
78	94	36	90	50	72	38
79	92	65	92	50	94	33
80	95	62	97	37	96	28
81	79	09	73	40	41	26
82	94	62	95	47	95	31
83	85	40	88	20	92	08
84	84	17	83	28	69	43
85	71	07	68	23	79	12

TABLE 4.2 (cont.)

	First Grade		Second Grade		Fourth Grade	
Item	<u>D</u>	<u>C</u>	<u>D</u>	<u>C</u>	<u>D</u>	<u>C</u>
86	89	56	85	39	73	25
87	78	23	90	67	88	36
88	92	49	96	12	99	17
89	76	21	77	26	60	22
90	80	30	73	38	61	36
91	86	40	93	46	94	28
92	86	42	89	42	93	41
93	90	50	92	43	89	37
94	92	45	97	36	96	13
95	94	33	92	42	90	21
96	96	20	95	16	95	31
97	92	36	92	26	94	21
98	94	39	96	34	95	19
99	94	22	93	24	90	35
100	90	44	86	39	88	35

patterns of response from subjects in different grades. Seventeen items increased in difficulty (i.e., decreased in percentage choosing the keyed answer) with increasing grade, at the same time that the item-test correlations for these items increased, with only two exceptions. This indicates that for these items the tendency for only children with high achievement motivation to choose the right alternative increased with age. It may be that extraneous factors related to the visual and/or content attractiveness of the motivated alternative accounts for the relatively high scores of first-graders.

For the first-graders, eight items had item-test correlations below .10, while the highest correlation was .67. For second-graders, 10 items had item-test correlations below .10, and the highest was .54. For the fourth-grade group, 13 items had item-test correlations of .10 or less, while the highest correlation was .60.

In considering these data, it should be kept in mind that the 100 items comprising the test were selected from a set of 200 items on the basis of item analysis for four-year-old Head Start children. Not only did these children differ from the elementary school children in age, but also they came from homes of a generally lower socioeconomic status. That some items are unsuitable for older children was to be expected.

The mean scores for these groups of elementary school children can be compared with a mean of 58.64 for 212 Hawaiian Head Start children, who were administered the same 100 items in an individual form of the test, before the format was randomized, in the fall of 1968. These children had a mean age of around four years and three months, with a range of close to 12 months.

For the 122 Head Start children who were in the national evaluation sample in the fall of 1968 and whose 100-item pre-tests were rescored on the basis of a selected 55 items, there was evidence of a low but positive correlation with age.

It seems probable that with the omission of certain items the test scores would increase to at least a small extent from the first to the fourth grade. On the other hand, there is ample reason to believe that motivation to achieve does not increase indefinitely with age. Indeed, for some persons it certainly decreases as they grow older.

Data for the 75-item randomized Individual Form of Gumpgookies for Ten Ethnic-cultural Groups

One of the purposes of this study was to obtain data relevant to possible effects of ethnic and cultural differences upon Gumpgookies scores. To this end, the first 100-item form of the test, which had been based upon preliminary analyses of a 200-item form given in two administrations, was revised. The test was shortened to 75 items; the wording of many items was changed, primarily to reduce the verbal difficulty; the illustrations were improved, usually in an effort to avoid a tendency to respond to particular features of the figures, such as uplifted arms; the positions of the figures on the pages were altered to include up and down, lower-left and upper-right, and upper-left and lower-right, as well as left and right illustrations; the order in which the illustrations were described by the examiner and the positions of the keyed answers were randomized.

The test was given to 10 groups in the spring of 1969. To the extent possible, these groups consisted of children between the ages

of four and five. Also, where possible, they consisted of Head Start children. In the case of Catholic, Mormon, and Jewish children, however, it proved impossible to locate sufficiently large samples in Head Start classes. Although interest lay in the possibility of differences associated with religious background, comparisons are confounded by the differences of these groups from others in socioeconomic background as well. An additional confounding factor is that these groups were concentrated in cities.

Two hundred and twenty-five Mormon children were tested in Utah. About one-half of the sample lived in and around Brigham Young University, with some enrolled in experimental classes at the University. The rest of the children came from preschools and special reading classes in Salt Lake City. They can be considered as having come from predominantly middle-income homes, with many parents being faculty or students at Brigham Young University.

Two hundred Catholic children were tested in New York City. Part of the sample was drawn from experimental classes associated with a college, some attending Day-Care Centers and some attending public schools. Part of the sample came from low-income areas and the rest were from middle-income areas.

One hundred and ninety-seven Jewish children were tested in New York City. They were attending preschools sponsored by Jewish religious and cultural organizations and not Head Start. Their socioeconomic level was predominantly middle-class.

Ninety Puerto Rican children were found in various Head Start Centers and preschools in New York City. Most of them were bilingual and met the poverty criteria set by O.E.O.

Two hundred and two Negro children were tested in Los Angeles, California. Approximately half of the children were enrolled in Head Start and thus met the low-income criterion set by O.E.O. The other half were obtained from Children's Centers. The coordinator for this sample regarded the parents of this latter group as upward-mobile.

Thirty-six White children were tested in rural areas in Oregon. Most of the children were attending Head Start classes and met the low-income criterion set by O.E.O.

Forty-seven Hawaiian children were tested in various Head Start Centers in Oahu. For the purposes of certain analyses, 223 more children were added later, all of them also enrolled in Head Start classes in Hawaii.

Thirty-nine children of Oriental extraction were tested in San Francisco and Los Angeles. The majority were attending two state-sponsored preschools in the Chinatown section of San Francisco. They were bilingual and from low-income homes. The children from Los Angeles were attending Head Start and met the O.E.O. low-income criterion.

One hundred and fourteen Mexican-American children were tested in Texas. Approximately half of the sample were children attending Day-Care Centers in San Antonio, while the rest were attending Day-Care Centers in several small towns in rural areas. All of the children were bilingual and met the O.E.O. low-income criterion. Most of them were scheduled to attend the Summer Head Start classes.

Two hundred and thirty-four American-Indian children were tested in Montana and Oregon. All of the children lived on reservations and rural areas and met the criteria for admission to Head Start, although some were not currently enrolled in Head Start. They all can be described as meeting the O.E.O. low-income criterion.

It had been the hope, and indeed the expectation, that the test could be given to some children in American Samoa, in view of some hints from various sources that children in that cultural environment are not particularly encouraged to excel or to be any different from other children. Although preliminary arrangements were made, permission to give the test was not given.

Table 4.3 contains the means (in order of size), standard deviations, measures of skewness and kurtosis, KR-20 reliability estimates, ranges, and numbers of cases. That the Mormon, Catholic, and Jewish children have the highest means is consistent both with the hypothesis that these religious groups stress motivation to achieve as well as with the hypothesis that socioeconomic status of families is correlated with childrens' scores on the test. It should be recognized that other comparisons as well must be made with due caution, both because of small N's in some instances and because the sampling for all of the groups can make no pretense to being nationwide nor representative.

The relation of scores on the test for these groups of younger children to age is also of some interest. Since this study was undertaken in the conviction that motivation is learned or at least can be changed, a positive relation with age was expected. Although this relation probably is not high even over a very wide age range, since some persons never develop much motivation to achieve, within the relatively narrow age ranges of the 10 subsamples in this phase of this study positive but low correlations were expected.

Tables 4.4 and 4.5 show, respectively for boys and girls, the mean ages and Gumpgookies scores of the subsamples and the correlation coefficients between age and test score. For the total samples of boys

TABLE 4.3

Means, Standard Deviations, and Other Statistics
for the 75-Item Gumpgookies Test, Ten Subsamples

	Mormon	Catholic	Jewish	Puerto Rican	Negro Urban	White Rural	Hawaiian	Oriental West Coast	Mexican-American	American Indian
Mean	1 61.17	2 58.86	3 58.74	4 56.33	5 55.90	6 54.25	7 53.00	8 52.46	9 52.19	10 50.25
S.E.	.56	.64	.72	.95	.53	1.53	1.19	1.32	.85	.64
Standard Deviation	8.38	9.09	10.05	9.05	7.57	9.18	8.15	8.27	9.13	9.72
S.E.	.42	.43	.60	.52	.42	1.15	.68	.81	.56	.40
Skewness	-.86	-.60	-.83	-.29	-.65	-.79	.27	-.15	-.19	.14
S.E.	.16	.17	.17	.25	.17	.39	.35	.38	.23	.16
Kurtosis	.25	-.20	.79	-.80	.55	.27	-.71	-.49	-.29	-.40
S.E.	.32	.34	.34	.50	.34	.77	.68	.74	.45	.32
KR-20	.86	.87	.89	.85	.77	.85	.79	.79	.83	.84
Range										
Low	35	33	18	37	29	32	38	35	25	25
High	75	74	75	73	70	70	70	70	70	74
N	225	200	197	90	202	36	47	39	114	234

TABLE 4.4

Means and Standard Deviations of Age and
Gumpgookies Score and the Correlation Between
 Them for Males in Ten Subsamples and the Total Group of Males

Sample	Age (in months) (Mean and Standard Deviation)	<u>Gumpgookies</u> Score (Mean and Standard Deviation)	N	Pearson r
1) Mormon	59.42 6.11	59.96 8.56	106	.30
2) Catholic	58.61 4.60	58.10 9.14	102	.27
3) Jewish	56.48 5.29	56.40 9.75	83	.09
4) Puerto Rican	59.87 4.82	55.96 8.65	52	.39
5) Negro Urban	58.01 4.32	57.03 7.11	106	.11
6) White Rural	55.33 7.26	55.04 9.14	24	.18
7) Hawaiian	58.81 3.81	54.04 7.86	26	.03
8) Oriental West Coast	60.39 4.43	53.61 9.24	18	.20
9) Mexican-American	61.79 5.96	50.96 9.09	53	.36
10) American Indian	56.02 5.63	49.56 9.12	122	.22
Total Male Sample	58.21 5.54	55.43 9.39	692	.23

TABLE 4.5

Means and Standard Deviations of Age and Gumpgookies
Score and the Correlation Between Them for Females in
Ten Subsamples and the Total Group of Females

Sample	Age (in months) (Mean and Stan- dard Deviation)	<u>Gumpgookies</u> Score (Mean and Stan- dard Deviation)	N	Pearson r
1) Mormon	59.62 5.30	61.86 8.43	108	.24
2) Catholic	58.54 4.86	59.51 9.05	96	.25
3) Jewish	57.05 4.80	60.83 9.73	111	.41
4) Puerto Rican	59.05 5.49	56.84 9.84	38	.44
5) Negro Urban	57.93 7.40	54.48 8.00	96	.12
6) White Rural	56.25 5.21	52.67 9.52	12	.31
7) Hawaiian	57.43 3.84	51.95 8.46	21	.23
8) Oriental West Coast	58.10 4.77	51.52 7.43	21	-.26
9) Mexican-American	61.88 5.58	52.83 9.17	60	.26
10) American Indian	56.64 5.84	50.92 10.32	112	.26
Total Female Sample	58.30 5.76	56.61 10.02	675	.24

and girls, mean ages, mean test scores, and correlations between age and score are highly similar. The mean for boys and girls combined is 56.01, or 74.68% of the total number of items. For six subsamples, girls' mean test scores exceed those of boys, but for three of these samples the girls were slightly older on the average. More elaborate statistical analyses could have been done. Inspection of the data, however, seems quite sufficient to warrant the general conclusions that, for children averaging slightly under five years of age and in general ranging between slightly over four to about five and three-fourths years of age, (a) sex differences in total score are relatively small; (b) sex differences in correlation of test score with age are relatively small; and (c) the correlation between test score and age is in the positive direction and of about the order of magnitude that was expected. Further exploration of the higher mean performance of the Jewish girls in comparison with Jewish boys might possibly be worthwhile, though, as noted before, the average age of the girls was a bit higher and the standard deviation lower. The finding that mean total scores do not differ appreciably for boys and girls does not preclude the likelihood that particular items may behave differently--i.e., have different difficulty values, item-test correlations, factorial composition, and predictive significance--for the two sexes. Of course, strictly speaking, the inspection of data in Tables 4.4 and 4.5 has not ruled out the possibility that the total test score could have different factorial composition and predictive significance for boys and girls.

Table 4.6 shows for the Gumpgookies items the difficulty values (percentages passing) and the point-biserial correlation coefficients of the items with the total score (75-item form) for each of the

TABLE 4.6

Difficulty Values (Percentages Passing)
and Item-Test Correlations (Point-Biserials)
of Items in the 75-Item Gumpgookies for Ten Subsamples

Item	1 Mormon		2 Catholic		3 Jewish		4 Puerto Rican		5 Negro Urban		6 White Rural		7 Hawaiian		8 Oriental-West Coast		9 Mexican-American		10 American Indian	
	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C
1	93	17	89	27	87	30	81	35	90	18	86	42	96	22	90	11	85	19	80	23
2	84	26	81	26	79	21	70	31	79	26	78	31	83	24	77	44	65	30	66	12
3	90	20	84	28	81	31	76	47	82	33	64	01	89	32	67	47	74	39	66	28
4	80	29	75	46	78	35	61	13	80	21	72	59	79	36	80	35	68	35	60	15
5	89	27	88	36	87	39	76	30	73	29	92	25	66	47	77	19	73	27	61	16
6	82	38	80	31	71	42	81	35	70	32	64	35	62	39	56	37	77	29	69	30
7	89	13	73	16	68	23	73	22	65	17	69	-03	47	03	64	12	65	27	66	18
8	92	20	86	42	90	41	84	47	82	32	86	25	85	21	77	26	83	25	67	30
9	76	36	72	23	71	34	64	38	67	22	67	42	62	37	82	12	63	24	71	33
10	79	39	69	20	75	21	67	21	73	19	58	06	77	01	64	01	66	21	73	09
11	83	34	83	29	82	49	84	24	82	32	94	-11	87	25	72	45	81	18	63	15
12	91	43	87	41	93	39	86	23	90	35	92	24	83	31	74	55	76	21	75	24
13	82	38	84	27	84	35	83	17	78	28	83	25	87	06	87	18	90	22	81	25
14	63	39	46	31	56	44	40	28	62	21	53	43	57	-03	46	30	47	33	47	21
15	70	28	42	31	56	39	44	28	65	17	36	16	55	34	41	20	47	15	39	28
16	88	23	83	20	84	30	69	27	75	32	75	12	79	03	80	28	65	41	66	24
17	86	24	91	31	90	26	86	31	81	24	83	04	77	41	72	34	79	42	79	31
18	82	15	79	22	76	19	86	15	74	16	72	35	83	15	85	-11	75	09	80	30
19	78	36	67	37	66	44	62	27	68	19	58	34	60	14	77	15	58	07	62	26

TABLE 4.6 (cont.)

Item	1 Mormon		2 Catholic		3 Jewish		4 Puerto Rican		5 Negro Urban		6 White Rural		7 Hawaiian		8 Oriental-West Coast		9 Mexican-American		10 American Indian	
	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C
20	75	33	78	22	76	40	74	30	69	21	67	24	64	25	77	29	57	18	62	32
21	85	20	77	31	69	30	76	25	81	16	75	41	77	40	67	45	72	22	68	24
22	88	21	84	25	90	22	79	25	86	21	81	29	75	18	69	11	64	14	73	30
23	84	42	82	35	82	38	82	30	81	30	81	39	57	58	67	53	75	37	64	43
24	81	46	78	36	78	34	72	27	76	25	67	43	81	19	77	10	62	22	65	34
25	81	35	78	19	70	17	74	02	70	14	67	28	75	17	51	-00	60	01	61	01
26	85	32	81	43	77	56	81	27	77	23	72	75	70	37	74	-11	70	36	67	45
27	82	22	81	23	76	15	80	32	77	28	83	-04	79	41	72	15	72	21	74	21
28	67	25	65	16	56	24	63	20	60	25	44	-16	53	-24	46	19	55	24	50	10
29	58	28	73	31	70	32	66	23	68	28	61	51	57	41	64	06	62	19	57	39
30	76	23	83	25	69	37	74	19	75	19	72	34	70	49	64	-08	69	41	65	34
31	81	49	81	39	79	41	74	26	75	21	78	41	68	47	56	29	80	33	65	39
32	85	31	84	39	76	34	79	54	80	14	69	31	81	09	82	38	75	36	68	20
33	92	05	88	13	93	17	82	27	82	01	78	-09	72	02	82	-10	74	33	74	38
34	85	41	78	54	77	45	69	32	76	43	72	68	53	40	62	35	68	30	61	38
35	85	31	76	33	77	42	73	31	75	35	75	53	70	26	64	37	64	19	69	44
36	84	24	89	42	89	47	91	37	75	28	81	22	68	27	80	17	82	19	73	18
37	75	23	78	22	71	28	78	43	73	25	83	42	89	20	74	13	77	34	69	25
38	93	21	91	23	86	23	89	22	79	22	86	28	70	41	72	21	76	25	80	38
39	61	13	58	14	62	27	58	03	68	13	67	19	68	00	69	25	59	16	70	12
40	77	48	78	30	79	42	80	33	75	25	69	19	47	14	54	15	75	28	70	42
41	66	02	67	18	65	13	61	29	59	20	58	-15	60	-10	67	-18	61	17	61	18
42	78	29	72	29	72	36	60	22	66	22	72	40	64	-04	46	-10	60	24	62	09
43	79	22	82	16	82	31	72	27	66	28	67	03	72	35	72	30	74	31	68	45
44	91	35	87	29	86	34	90	36	77	27	89	48	87	39	80	41	81	43	78	45
45	84	36	83	51	83	44	87	37	75	29	69	58	68	50	74	26	75	37	75	37
46	75	31	61	23	62	29	68	27	66	12	44	38	55	01	62	-06	61	17	59	21

TABLE 4.6 (cont.)

Item	1 Mormon		2 Catholic		3 Jewish		4 Puerto Rican		5 Negro Urban		6 White Rural		7 Hawaiian		8 Oriental-West Coast		9 Mexican-American		10 American Indian	
	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C
47	66	20	70	17	70	11	64	07	59	-00	36	14	75	14	64	09	57	12	50	-12
48	57	22	63	12	64	19	52	08	66	22	50	11	45	-04	62	41	49	09	48	25
49	87	25	80	20	78	44	77	31	79	26	78	38	83	07	69	24	72	33	76	20
50	91	50	94	37	92	34	86	40	87	25	83	42	89	11	74	47	75	39	73	43
51	91	36	85	42	94	30	82	41	74	34	81	27	75	33	59	43	61	21	67	38
52	85	39	82	29	81	50	86	34	70	34	72	33	77	41	69	32	75	34	72	29
53	78	46	76	51	70	49	73	29	75	25	64	30	68	25	80	10	78	39	77	36
54	89	34	85	44	85	40	82	32	74	31	75	13	70	31	74	23	74	21	90	32
55	76	38	77	35	78	38	74	31	73	14	69	12	66	06	69	38	61	38	60	48
56	84	58	85	41	89	41	81	37	75	29	64	57	64	33	77	24	75	35	71	45
57	88	27	79	43	88	32	86	37	77	30	89	24	87	24	64	31	75	30	70	33
58	77	48	69	42	70	42	70	30	73	21	64	18	64	23	69	50	64	16	55	32
59	85	32	88	50	85	44	86	55	79	40	61	59	77	35	82	30	73	46	68	31
60	77	02	80	23	84	27	78	10	72	32	67	20	64	12	80	31	73	19	65	11
61	87	30	84	36	79	38	79	35	80	35	89	56	55	52	74	64	71	25	69	37
62	77	27	76	38	75	26	78	46	75	19	75	37	68	23	87	31	69	43	73	27
63	89	27	86	27	90	31	84	37	83	14	92	-06	75	42	69	35	72	37	72	35
64	88	30	90	27	84	31	84	23	82	16	75	39	81	39	77	-01	89	35	80	30
65	71	22	78	27	74	38	74	19	72	21	83	-27	57	15	69	22	68	36	52	15
66	83	37	81	47	83	40	74	40	70	27	75	39	51	44	59	04	68	31	64	41
67	84	32	84	28	89	31	80	30	73	29	72	14	66	32	80	17	76	38	79	38
68	82	44	77	44	81	45	70	51	75	37	78	52	49	45	59	56	59	40	59	47
69	88	30	83	37	85	41	80	57	81	17	75	62	79	36	67	53	74	34	66	31
70	87	24	87	34	80	46	79	35	79	22	81	40	55	16	77	16	76	36	66	41

TABLE 4.6 (cont.)

Item	1 Mormon		2 Catholic		3 Jewish		4 Puerto Rican		5 Negro Urban		6 White Rural		7 Hawaiian		8 Oriental-West Coast		9 Mexican-American		10 American Indian	
	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C	D	C
71	88	21	83	25	86	32	73	33	70	09	89	06	89	07	90	-09	68	15	79	19
72	83	39	76	32	73	38	69	33	66	25	69	43	68	15	59	62	69	22	60	27
73	79	23	79	12	76	30	78	25	71	07	75	25	87	22	56	34	74	25	74	33
74	84	22	82	26	79	18	79	03	78	03	75	01	87	32	74	40	75	17	77	25
75	84	30	74	29	75	26	68	02	77	37	58	20	66	40	67	05	58	27	58	14

10 subsamples. Inspection of this table reveals that no item had a difficulty below the chance value (50%) for all samples, but one (# 14) had a value below 50% for five samples and one (# 15) such a value for six samples. No item correlated negatively with the total test score for more than three samples. Without presentation of detailed discussion of these and related findings for each item with respect to each subsample, the investigators commit themselves to careful inspection of these data in undertaking the next revision of the test.

Before this task has been accomplished, however, the reader may find some interest in the correlations among the item difficulty values and among the item-test correlation coefficients for the several subsamples. These are presented in Tables 4.7 and 4.8, respectively. Although certain groupings of the subsamples--whether on religious, racial, ethnic, socioeconomic, or a rural-urban dimension or a confusing concatenation of these influences may still depend upon one's biases--will be apparent from the foregoing tables, a more systematic approach is available via the route of factor analysis. Accordingly, the two matrices based upon correlations among difficulty indices and among item-test point-biserial correlation coefficients were each factor-analyzed to three factors by a principal-axes solution with orthogonal rotation.

Table 4.9 presents the orthogonally rotated loadings for the factors based upon item difficulty indices. Recall that the subsamples are listed in order of total mean score on the test, from high to low. The highest loading for each subsample was first identified. Factor 1 was then found to be associated with subgroups (in order of magnitude

TABLE 4.7

Correlations (Pearson r 's) Among Difficulty
Indices of Gumpgookies Items for Ten Subsamples

	Mormon (N=225)	Catholic (N=200)	Jewish (N=197)	Puerto Rican (N=90)	Negro (Urban) (N=202)	White (Rural) (N=36)	Hawaiian (N=47)	Oriental (W.C.) (N=39)	Mex.-American (N=114)	Am. Indian (N=234)
Mormon (N=225)	--	.77	.77	.72	.72	.67	.48	.39	.62	.62
Catholic (N=200)	.77	--	.87	.88	.70	.74	.47	.55	.77	.68
Jewish (N=197)	.77	.87	--	.78	.71	.71	.47	.51	.66	.63
Puerto Rican (N=90)	.72	.88	.78	--	.65	.68	.46	.49	.82	.73
Negro (Urban) (N=202)	.72	.70	.71	.65	--	.68	.53	.47	.63	.53
White (Rural) (N=36)	.67	.74	.71	.68	.68	--	.45	.43	.65	.59
Hawaiian (N=47)	.48	.47	.47	.46	.53	.45	--	.46	.52	.54
Oriental (W.C.) (N=39)	.39	.55	.51	.49	.47	.43	.46	--	.53	.59
Mex.-American (N=114)	.62	.77	.66	.82	.63	.65	.52	.53	--	.73
Am. Indian (N=234)	.62	.68	.63	.73	.53	.59	.54	.59	.73	--

TABLE 4.8

Correlations (Pearson r 's) Among Gumpgookies
Item-Test Correlation Coefficients for Ten Subsamples

	Mormon (N=225)	Catholic (N=200)	Jewish (N=197)	Puerto Rican (N=90)	Negro (Urban) (N=202)	White (Rural) (N=36)	Hawaiian (N=47)	Oriental (W.C.) (N=39)	Mex.-American (N=114)	Am. Indian (N=234)
Mormon (N=225)	--	.57	.57	.26	.34	.39	.26	.24	.21	.40
Catholic (N=200)	.57	--	.66	.50	.49	.54	.44	.31	.35	.40
Jewish (N=197)	.57	.66	--	.44	.47	.50	.29	.26	.36	.41
Puerto Rican (N=90)	.26	.50	.44	--	.33	.31	.23	.36	.52	.46
Negro (Urban) (N=202)	.34	.49	.47	.33	--	.30	.41	.32	.27	.23
White (Rural) (N=36)	.39	.54	.50	.31	.30	--	.33	.14	.26	.33
Hawaiian (N=47)	.26	.44	.29	.23	.41	.33	--	.28	.32	.42
Oriental (W.C.) (N=39)	.24	.31	.26	.36	.32	.14	.28	--	.23	.28
Mex.-American (N=114)	.21	.35	.36	.52	.27	.26	.32	.23	--	.44
Am. Indian (N=234)	.40	.40	.41	.46	.23	.33	.42	.28	.44	--

TABLE 4.9

Orthogonally Transformed Factor Loadings
Based Upon Factor Analysis of the Item Difficulty
Values for Ten Subsamples (75-Item Form of Gumpgookies)

	Factors *		
	<u>1</u>	<u>2</u>	<u>3</u>
1. Mormon	83	18	27
2. Catholic	82	47	10
3. Jewish	82	34	18
4. Puerto Rican	76	54	04
5. Negro (Urban)	74	16	46
6. White (Rural)	78	25	23
7. Hawaiian	25	32	86
8. Oriental (West Coast)	15	80	33
9. Mexican-American	61	62	14
10. American Indian	46	73	21

* Decimal points have been omitted.

of factor loadings) 1, 2, 3, 6, 4, 5 (Mormon, Jewish, Catholic, White-rural, Negro-urban, and Puerto Rican-urban). Factor 2 was associated with groups 8, 9, and 10 (Oriental West Coast, American Indian, and Mexican-American). However, the loading for the Mexican-American group was almost as high for factor 1 as for factor 2. The Hawaiians (group 7) stood alone on factor 3.

In the factor analysis of the relationships among the item-test correlation coefficients for the various groups (Table 4.10), factor 1 is associated, in order, with subsamples 3, 1, 2, and 6 (Jewish, Mormon, Catholic, and White-rural). Factor 2 is associated, in order, with Mexican-American, Puerto Rican, and American-Indian subsamples. Factor 3 has its highest loadings, in order, for the Oriental West Coast, urban Negro and Hawaiian children in the subsamples.

Different readers may be tempted to speculate at length, and to different ends, about the foregoing results. One can scarcely escape the finding that groups identified as Jews, Catholics, Mormons, and rural Whites are found on both a common difficulty and a common item-test correlation factor. Mexican-Americans and American Indians go together on a factor in each of the analyses. Puerto Ricans appear on the same item-correlation factor; but, although they have a high loading on the same factor for item difficulty, they have a still higher loading on factor 1. The urban Negroes, joining the Mormons, Catholics, Jews, rural Whites and Puerto Ricans on the first item-difficulty factor, go with the Oriental West-Coasters and Hawaiians on the third item-difficulty factor.

No definite conclusions as to subsample differences as reflected in these analyses are here advanced. Various combinations of differences

TABLE 4.10

Orthogonally Transformed Factor Loadings Based
 Upon Factor Analysis of the Item-Test Correlation
 Coefficients for Ten Subsamples (75-Item Form of Gumpgookies)

	Factors *		
	<u>1</u>	<u>2</u>	<u>3</u>
1. Mormon	76	08	18
2. Catholic	74	28	33
3. Jewish	77	26	21
4. Puerto Rican	22	72	26
5. Negro (Urban)	43	06	66
6. White (Rural)	73	25	-01
7. Hawaiian	27	32	49
8. Oriental (West Coast)	00	20	82
9. Mexican-American	11	83	12
10. American Indian	31	69	14

*Decimal points have been omitted.

in socioeconomic levels, religious backgrounds, ethnic effects, cultural influences, and bilingualism versus monolingualism in the home might be brought forth as explanatory hypotheses. Some may even suspect the hodge-podge popularly referred to as the "white, middle-class" background of the perpetrators of the Gumpgookies test.

Nonetheless, if members of subsamples who do less well on the test aspire to what they perceive as advantages of subsamples who do better on the test--in whatever guise--then measurement of motivation to achieve and investigation of means of increasing it should be palatable.

CHAPTER 5

DATA PERTAINING TO THE VALIDITY AND RELIABILITY OF GUMPGOOKIES

Relation of Gumpgookies to External Criterion Measures

In addition to study of the internal validity of Gumpgookies by means of factor analytical techniques, as described in Chapter 6, several attempts have been made to relate the Gumpgookies test to other independent measures of motivation to achieve and related behaviors.

The first effort involved a specific arranged task on the performance of which two observers each rated the children on seven hypothesized motivational dimensions. The task consisted of putting together a puzzle during a 15-minute period. The children were each handed a puzzle but worked at a table with two to five others. The teacher took charge of the group, framing questions and allowing the children also to ask questions. She was available for help and worked with children who were experiencing consistent failure. Specific instructions were given as to what she could say. She was instructed to praise or encourage each child approximately three times during the session. Easier puzzles were available for the child who found the first puzzle too difficult and more complicated ones for the child who finished early. After 12 minutes, each child was given the option of going out to play or finishing his puzzle.

The observers watched for signs of and rated behavior representing the following characteristics:

1. eagerness to begin
2. inattention

3. seeking or responding to praise and showing signs of pride
4. accepting or seeking help, or not needing it
5. comments of interest or curiosity questions
6. carefulness
7. persistence

Rating was done in three five-minute intervals. Eagerness to begin and persistence were rated on a scale, while evidence of other dimensions was recorded each time upon occurrence. After the observation, these characteristics were assigned ratings depending on number of occurrences. The three ratings for each characteristic were averaged over the 15 minutes, and each subject was given a total score.

The Pearson r between the motivation ratings assigned by the two raters during the puzzle task (inter-rater reliability) was .83 ($N = 13$). The Pearson r between the average ratings assigned by the two raters and the Gumpgookies scores was .28 ($N = 13$), not significant at the .05 level. Although the ratings were apparently reliable, the behavior observed may have related to factors other than motivation. One weakness that had been anticipated was that the task used may have resulted in behavior that was too situation-specific or too time-specific. Thus a particular child's behavior may have varied from day to day or from one problem to another. Part of the difficulty, also, seemed to be attributable to the interaction of the teacher with particular children. She gave certain children much more attention, prodding some more than others. Therefore, some children were encouraged more than others to ask questions, seek help, verbalize, and persist on the task, thereby receiving higher ratings. The teacher's behavior (i.e., enthusiasm) changed from day to day. In addition, limited spread of the scores on the test can contribute to a low correlation coefficient. When other criterion measures were deemed more satisfactory, this method was not

further pursued. One suggestion, if the method were to be tried again, was to use a tape recorder or video tape to better assess behavior of the child and his interactions with the teacher.

Two different behavior rating scales were completed by two teachers within one classroom. The first was Zigler's Behavior Inventory, adapted by Project Head Start for use in 1966-67. The teachers rated each child on 50 items, using a four-point scale for each. Only 12 of these items (those deemed to be related to motivation) were analyzed. All were rated, however, with a view to obtaining more reliable and valid ratings of the factor seemingly related to motivation.

The regular classroom teacher's ratings yielded a rank-difference correlation of .48 (significant at the .05 level, $N = 16$) with the Gumpgookies scores. A motivation behavior rating scale composed by Ballif was also completed by the two teachers. The regular classroom teacher's motivation rankings correlated .58 with the Gumpgookies scores (significant at the .01 level). The part-time language teacher's rankings correlated .72 with Gumpgookies (significant at the .01 level). Both of these rating devices appeared to be good criterion measures. They were not used, however, as a still simpler method, perhaps ill-advisedly, in retrospect, was deemed feasible.

The same two teachers were asked to read a description of motivation and to rank the children with respect to motivation. The rankings made by the regular classroom teacher had a rank-difference correlation of .59 with Gumpgookies scores, significant at the .05 level. The rankings of the part-time language teacher correlated .22 with Gumpgookies (not significant). However, the rank-difference correlation between the two teachers' rankings, .62, was significant. Both teachers' rankings correlated with their ratings on the Zigler Behavior Inventory

at a significant level. It was felt this ranking procedure would be a fairly simple and accurate measure to use as a criterion.

The written definition of motivation and instructions for ranking children were presented to 14 additional Head Start teachers. Motivation was described as something within a person that incites him to action or effort in certain directions. The idea that motivation is both an attitude and a process, including ways of feeling as well as ways of reacting, was explained. Motivation was differentiated from the concepts of achievement and aptitude. Finally, examples of the motivated child were included. Teachers were instructed to begin by listing the child showing the highest motivation, then the one with the lowest, next the second highest, then the second lowest, and so on.

Rank-difference correlation coefficients between the ranks on the test and the ranks assigned by the teacher were computed. On the basis of a standard table of critical values for this type of coefficient, five of the 14 coefficients were beyond the .05 significance level, and 12 of the 14 coefficients were positive. In view of the somewhat more promising results with the Zigler and Ballif scales, further work based on them is indicated.

The Gumpgookies scores of the three most highly motivated and the three least motivated children, as designated by the teacher's rankings on motivation, were analyzed separately by means of the median test. The rankings were completed by 10 teachers, yielding a sample of 60 children. A Gumpgookies score of 57 was the median. Of the 30 students ranked most highly on motivation, 17 were above the median, 10 below the median, and three at the median. Of the students ranked lowest in motivation, 10 were above the median, 19 below the

median, and only one at the median. A chi-squared analysis of this breakdown yielded a value of 4.62, significant at the .05 level.

An analysis of variance computed on these same data yielded an F value of 13.07, significant at the .01 level (Table 5.1). A similar analysis with these data, plus data from five additional classes collected later in the year, yielded an F significant at the .05 level (Table 5.2).

Some additional data are available with the group form of the Gumpgookies for first-, second-, and fourth-graders. Each teacher was asked to rank the top 15% and the bottom 15% of her class with respect to motivation with the same instructions as had been used before. The differences in Gumpgookies scores between highest and lowest ranked children for each grade were subjected to analysis of variance procedures. These data are presented in Table 5.3. The difference for the first grade was significant at the .05 level; that for the second grade was also significant at the .05 level; and the difference for the fourth grade, while favoring the group ranked high on motivation, was not significant.

Relation of Gumpgookies to the Children's Self-Concept Index

In the course of the development of Gumpgookies--and long after the format had been decided upon and the early editions of the test had been applied to a large number of subjects--the Children's Self-Concept Index came to the attention of the staff. This is an instrument adapted by the Westinghouse Learning Corporation from a test originally developed by Joseph H. Meyerowitz, who had been interested in an index of "self-derogation." The test was used in the National Impact Study

TABLE 5.1

Analysis of Variance of Gumpgookies Scores for the Three Highest and Three Lowest Ranked Students on Motivation, for Ten Classes

Group	High Motivation	Low Motivation
N	30	30
Mean	62.20	54.70
S. D.	9.05	6.87

Table of Variances

Source	Sums of Squares	D. F.	Mean Squares	F Ratio	p
Total	4586.84	59			
Between Groups	843.75	1	843.75	13.07	<.01
Within Groups	3743.09	58	64.54		

TABLE 5.2

Analysis of Variance of Gumpgookies Scores for the Three Highest and Three Lowest Ranked Students on Motivation, for Fifteen Classes

Group	High Motivation	Low Motivation
N	45	45
Mean	62.62	53.87
S. D.	9.52	6.36

Table of Variances

Source	Sums of Squares	D. F.	Mean Squares	F Ratio	p
Total	7492.61	89			
Between Groups	1724.84	1	1724.84	26.32	<.01
Within Groups	5767.77	88	65.54		

TABLE 5.3

Analyses of Variance of Gumpbookies Scores for the Fifteen Per Cent Highest and Fifteen Per Cent Lowest Ranked Students on Motivation

Group	First Grade		Second Grade		Fourth Grade	
	High	Low	High	Low	High	Low
Mean	84.40	75.68	87.27	81.00	82.03	80.24
S. D.	10.05	17.92	6.11	12.04	7.96	8.15
N	25	25	30	30	29	29

Sources of Variance

	First Grade	D. F.	Second Grade	D. F.	Fourth Grade	D. F.
Total S.S.	11077.89	49	5874.92	59	3680.89	57
Between Groups S.S.	950.48	1	589.07	1	46.62	1
Within Groups S.S.	10127.41	48	5285.99	58	3634.27	56
Between Groups M.S.	950.48		589.07		46.62	
Within Groups M.S.	210.99		91.14		64.90	
F Ratio	4.50		6.46		.72	
P Level	< .05		< .05		N.S.	

for Project Head Start conducted by Westinghouse. Its "internal consistency reliability" (KR-20?) has been reported to be .80, based on 100 second-grade pupils; test-retest reliability after two weeks was .66 (for the same sample?).

Each of 26 items contains a pair of "stick" figures, one a "balloon-child" and one a "flag-child." The balloon-child always appears in the left-hand position, and it is always described first.

Although the emphases of the items in this test and Gumpgookies do differ, there was sufficient similarity that a comparison of results of administering both tests to the same sample seemed worthwhile. In addition, the problem of right-left position preference (or primacy versus recency) had been encountered with the original format of Gumpgookies, in which the examiner uniformly described the left-hand figure first. Because of the format of the Self-Concept test, it was thought that the same problem might occur, even for second-grade children. The problem was regarded as not so likely to be evident as in the case of younger children, however. The test had been included in the Westinghouse study for groups of first-, second-, and third-graders.

The Self-Concept test and a 100-item, randomized form of Gumpgookies were administered to a sample of 104 children in the second grade in Hawaii schools in the spring of 1969. The correlation between the two tests was .43.

For full exploration of the question of position preference, it would have been desirable to set the items up for an individual form of the test to be administered to four-year-olds. Unfortunately, resources for the study did not permit the amount of work that this would have required.

It was possible, however, to factor the Self-Concept test. Two solutions were obtained, one with six factors and one with two factors. As had been found with factor analyses of other binary data covered in this report, the last eigenvalues exceeded unity; but, because of the difficulty of identifying as many as six factors with binary data and a small number of variables, no further solutions were obtained.

Both oblique and orthogonal rotations were obtained, the results for the two solutions being quite similar. Table 5.4 presents the obliquely rotated six-factor matrix, Table 5.5 the obliquely rotated two-factor matrix. In the two-factor solution, for which the majority of the items had their highest loadings on the first factor, there is little suggestion of position factors. In the six-factor solution, the five items with highest loadings on the second factor all have answers in the right-hand position. Thus, also, all of the answers indicate preference for the balloon-man, and all of the items had the correct answer presented last. For the first factor, seven of the highly loaded items had answers to the left, three to the right. The numbers of items for the other factors are too small to be indicative. For the two-factor solution, the highly loaded items on each factor are about equally divided between left and right. The results, while somewhat inconclusive, appear to indicate that if a position-preference problem is present, at least the test is measuring something else in addition when it is given to second-graders. From the results previously reported for the Gumpgookies test, a position preference problem, confounded with primacy and recency factors, very probably would be present were the Self-Concept test to be given to much younger children, say, three or four years of age, for whom the test would be

TABLE 5.4

Oblique (Oblimin) Factor Loading Matrix for a Six-Factor
 Solution for the Children's Self-Concept Index
 (104 Hawaiian Second-Grade Children, Spring, 1969)

Item Number	Factor Loadings					
	I	II	III	IV	V	VI
1	-.08	-.13	-.01	.43	.05	-.08
2	.38	.32	.14	.52	.18	.28
3	.27	.55	.44	-.34	.32	.15
4	.20	.26	.21	-.05	.71	.17
5	.44	.19	.70	-.03	.08	.12
6	.69	.23	.28	-.17	.21	.09
7	.56	.09	-.22	.24	-.03	.43
8	.26	.14	.47	-.11	-.49	.33
9	.43	.71	.43	.15	-.13	.35
10	.78	.14	.41	.06	.03	.32
11	.25	.13	.22	.00	.00	.80
12	.10	.60	.21	-.26	.15	-.16
13	.47	.47	.36	.04	.00	-.29
14	.72	.21	.61	.23	.17	.37
15	.64	.26	.31	.46	-.24	.53
16	.58	.17	.54	.06	.30	-.15
17	.21	.62	-.02	.23	-.15	.14
18	.54	.10	.41	.15	.25	.38
19	.50	.06	.58	-.46	-.03	.18
20	.20	.17	.67	.32	.09	.25
21	.65	.32	.34	-.22	.12	.00
22	.58	.22	.45	.55	-.30	.29
23	.00	.71	-.03	-.09	.27	-.06
24	.32	.11	.36	.21	.63	.07
25	.47	.09	.80	-.11	.29	.04
26	.61	.19	.23	.15	-.03	-.01

TABLE 5.5

Oblique (Oblimin) Factor Loading Matrix for a Two-Factor
 Solution for the Children's Self-Concept Index
 (104 Hawaiian Second-Grade Children, Spring, 1969)

Item Number	Factor Loadings	
	I	II
1	.01	-.19
2	.43	.11
3	.27	.69
4	.22	.47
5	.56	.34
6	.55	.37
7	.42	-.15
8	.38	.02
9	.49	.45
10	.75	.16
11	.37	-.01
12	.04	.64
13	.39	.49
14	.81	.25
15	.71	-.04
16	.58	.40
17	.16	.29
18	.61	.16
19	.50	.30
20	.47	.18
21	.52	.45
22	.70	-.04
23	-.09	.62
24	.40	.32
25	.60	.42
26	.53	.16

more difficult. Clearly, if a set of items is on the average as easy as this test was for the second-graders, factors extraneous to the content would be expected to play a very secondary role.

No attempt will be made here to interpret the factors in the Self-Concept test, except to note tentative interpretations of the factors in the two-factor solution. The first of these (which in the six-factor solution splits into three factors--the first, third, and reflected fourth) perhaps can be interpreted as related to self-confidence or a positive self-image. The second (which is similar to the second factor in the six-factor solution) may be related to social adjustment (or adjustments to other persons). Because of the small numbers of items on the second factor, this interpretation should be regarded as tenuous at best. Since some of the items in the Gumpgookies test attempt to measure similar factors, it is perhaps not surprising to find the correlation of .43 between the two tests.

The KR-20 reliability estimate for the Self-Concept test for the Hawaii sample was .87. Table 5.6 presents the difficulty values (percentage passing each item) and the item-test point-biserial correlation coefficients for the items of this test. The mean difficulty value is .91. Table 5.7 shows the frequency distribution. The mean score for the Hawaii second-graders on this 26-item test was 49.60. It is scored by crediting two points for an item correct and one point for an item missed. The possible range is thus 26 through 52 rather than 0 through 26. The mean reported in the Westinghouse study for a sample of 221 second-graders is 46.43, and for a sample of 48 third-graders it is only 46.83. Although both of these values are lower than those for the second-graders in Hawaii, the Westinghouse study distributions, too,

TABLE 5.6

Difficulty Values (Percentages Passing) and Item-Test Point-Biserial
Correlation Coefficients for the Items in the Children's Self-Concept Index
(Hawaiian Second-Graders, N = 104, Spring, 1969)

Item Number	Percentage Passing	Correlation with Total Score
1	.97	.00
2	.84	.45
3	.85	.50
4	.95	.35
5	.87	.58
6	.90	.59
7	.93	.33
8	.87	.36
9	.93	.60
10	.90	.67
11	.96	.33
12	.93	.29
13	.93	.49
14	.92	.74
15	.90	.61
16	.91	.60
17	.81	.31
18	.89	.57
19	.88	.52
20	.95	.44
21	.89	.59
22	.92	.60
23	.96	.20
24	.91	.45
25	.93	.60
26	.88	.51

TABLE 5.7

Frequency Distribution of Scores on
the Children's Self-Concept Index
(Hawaiian Second-Graders, N = 104, Spring, 1969)

Score	Frequency
52	36
51	24
50	13
49	11
48	4
47	4
46	2
45	1
44	1
43	0
42	1
41	2
40	2
39	1
38	0
37	0
36	0
35	1
34	1

Mean = 49.60

Standard Deviation = 3.58

Variance = 12.82

Skewness = -2.43

Kurtosis = 6.28

must have been markedly skewed. A child with a score of 47, for example, missed only five of the 26 items. Over half of the subjects missed no item or only one.

The full effects (upon estimates of test reliability and of item discrimination indices) of extreme difficulty values (either high or low), combined with extreme skewness of distributions, have not been explored. Nonetheless, extreme difficulty values and extreme skewness clearly do affect the possible range of values of both reliability estimates and point-biserial correlation coefficients. For example, consider an item such that 91% of the subjects pass. (Ninety-one is the mean percentage of the Hawaii sample passing the items of the Self-Concept test.) With the particular extremely skewed distribution for the Hawaiian second-grade total test scores, the range of possible point-biserial correlation coefficients has been calculated to be restricted to $-.12$, on the negative side, to $.86$, rather than -1.00 to 1.00 . The highest possible negative value of a split-half correlation coefficient--given the restriction that the distribution of total scores is as skewed as the one for the Hawaiian sample--is $-.05$, only slightly below zero rather than approaching -1.00 . Hence the estimates of both KR-20 reliability coefficients and point-biserial correlation coefficients appear to be rather seriously inflated.

The extent of the inflation cannot yet be pinpointed as applied to other specific skewed distributions. But the inapplicability of ordinary test statistics, or at least of their usual interpretations, seems clear. Hence the reliability estimates and the item-test coefficients for the Childrer's Self-Concept Index are undoubtedly inflated. In addition, and perhaps more important, attempts to measure mean gain on an instrument

with such a low ceiling for the age ranges of concern would seem very likely to reveal no significant gains.

Aside from the fact that the test is too easy to be optimally discriminating for perhaps the upper two-thirds or three-fourths of the distribution, it is quite possible that it would be useful for younger children. Although Meyerowitz has indicated that he was successful in giving a group form of the test to children as young as three, an attempt to duplicate this finding with Gumpgookies led to the conclusion that use of a group form was not feasible for four-year-olds, even when the groups consisted of only five children. In view of the similarity of the test formats, this experience would suggest that the Self-Concept test not be administered as a group test to children below six years of age.

A preliminary report on the data obtained in the Westinghouse study for the first three school grades confirms the relatively low ceiling of the test, although the average scores do not reach the mean of 24 items correct that was obtained for Hawaiian second-graders (who were tested in the spring of the year). The Hawaiian sample, it should be noted, was not composed solely of children who had previously been exposed to Head Start; hence it is likely that the socioeconomic levels of their families were higher than for the Westinghouse samples. Nevertheless, the low ceiling of the test would appear to limit its usefulness as a measure for reflecting change for children in elementary schools.

Relation of Gumpgookies to Other Tests

The fact that performance on the test is related to general intelligence and language proficiency is demonstrated by correlations obtained between Gumpgookies and several tests administered to the Hawaii Head Start sample during the 1968 national evaluation study. For this sample,

quite restricted in range of ages, Gumpgookies scores correlated positively (.11) but non-significantly with age. On the other hand, Gumpgookies correlated .24 ($p < .01$) with Stanford-Binet I.Q., .31 ($p < .01$) with Preschool Inventory scores, and .18 ($p < .05$) with the Animal House subtest of the WPPSI.

The correlation of .23 ($p < .01$) between Gumpgookies and the Psycholinguistic Age score of the ITPA may indicate that the scores on both tests are partially affected by language variables or that both are affected in part by "general intelligence." But it should be kept in mind that the ITPA, as well as Gumpgookies, correlates with IQ. For a sample of 137 Hawaiian children, the Gumpgookies - ITPA r was .23, the Gumpgookies - IQ r was .24, and the ITPA - IQ r was .45. The correlation of Gumpgookies and ITPA with IQ partialled out was .14. With a standard error of the corresponding z equal to .09, this partial correlation cannot be regarded as significantly different from zero.

Finally, data for first-, second-, and fourth-graders from Hawaii public schools on the Gumpgookies and several tests of academic achievement and aptitude, among them the Metropolitan Readiness Test, the California Reading Test, and the School and College Ability Test (SCAT), indicate that a low but significant correlation exists between the Gumpgookies test of achievement motivation and presumed measures of actual achievement in school.

In summary of the results presented in this section, there is evidence that Gumpgookies is measuring aspects of motivation to achieve. The position-preference, item-position, and primacy-recency influences mentioned elsewhere in this report, plus differences in socioeconomic, ethnic, and cultural variables for the various samples studied, do not

yet warrant conclusions about the relation of the Gumpgookies test with sex and/or ethnic-cultural differences. These independent efforts in the direction of concurrent validation of the test, coupled with the results of the several factor-analytical procedures given in Chapter 6, nevertheless are interpreted as warranting further efforts to understand motivation better, to measure it, and to enhance it in children at an early age.

Data Relevant to the Reliability of Gumpgookies

The only routinely available type of reliability estimate for Gumpgookies has been KR-20 (which Cronbach refers to as alpha), as presented in Chapter 4. For many separate groups and for different forms of the test, the estimates have been consistently close to .80 for the 75-item form and well above .80 for the 100-item form. If the KR-20 estimate is to be regarded as an index of internal consistency--as indeed seems to be appropriate--and if the test reflects several factors--as seems to be amply supported in Chapter 5--then the KR-20 index can be expected to provide an under-estimate of reliability as compared with an estimate for a single-factor test of comparable length.

Clearly an estimate of reliability based upon the correlation of different forms of the test administered at different times is to be preferred. In view of other problems in development of the test, no attempt has as yet been made to develop different forms. Nor is any completely satisfactory estimate of test-retest reliability available--if ever an estimate of this type of coefficient can be regarded as satisfactory. In one instance, however, 44 children who had taken the 100-item, non-randomized form were also given the 75-item, randomized form, after an interval of two to three weeks. Many of the items in

the shorter form had been revised, both with respect to wording and illustrations. The correlation between scores on the two forms was .66. The ages of the subjects ranged from 52 to 64 months, with a mean of 58 months. Although the number of cases was small, the obtained coefficient can be regarded as reasonably satisfactory considering the differences in the items and in the test format, along with a relatively narrow age range.

Two other measures of relationship between test and retest are available for sets of items in the original format. The Head Start children included in the 1968-69 evaluation samples took the 100-item, non-randomized form as a pre-test and a specially selected 55-item, non-randomized form as a post-test, some five months later. The reason for this shift from pre-test to post-test was that examiners from other Head Start Evaluation and Research Centers objected to the length of the 100-item pre-test. Later, the pre-tests were rescored on the basis of only the selected 55 items contained in the post-test. This means, of course, that the context of the items differed in the two instances. The scores for the 124 Hawaiian Head Start children having been most readily accessible, the correlation between the 55 items common to the pre-tests and post-tests for this sample was computed and found to be .35. This figure, although of some interest, is not a good estimate of test-retest reliability because the time interval was sufficiently long that differential changes in the children could be expected to have occurred.

A similar comparison was made of the scores of the 74 Hawaiian Head Start children in the combined experimental and comparison groups in the motivation curriculum study. In this case, the 100-item, non-randomized

form was used. The correlation was .48. Again, although this coefficient could be expected to have some relevance to reliability, it is not an estimate of test-retest reliability.

Although the two aforementioned coefficients are not reliability coefficients in any meaningful sense, the shift from .35 to .48 in going from the 55-item form to the 100-item form is very close to the shift from .35 to .49 that the Spearman-Brown prophecy formula would predict.

CHAPTER 6

FACTOR ANALYSES OF RANDOMIZED FORMS OF GUMPGOOKIES FOR FIRST-, SECOND-, AND FOURTH-GRADERS AND FOR TEN GROUPS OF PRESCHOOLERS

The concept of motivation, somewhat like that of general intelligence, perhaps can be useful at different levels in a hierarchical structure. Hence a dominant concern of the series of investigations reported herein has been to develop a general measure of motivation to achieve in school that will be consonant with a theoretical base. At the same time, the score or scores resulting from the instrument should make sense in terms of empirical relationships with such variables as measures of IQ; other tests that might be expected to have a correlation with the motivation measure; tests purported to get at actual achievement; other persons' appraisals of motivation to achieve as distinct from aptitude and actual achievement; and, finally, such variables as age, sex, and the pervasive multiplicity of elements that have been blanketed under the term "ethnic-cultural" groups.

Previous chapters have presented evidence that a total score on the Gumpgookies test can be obtained, even for four-year-old children from impoverished environments, that has properties in line with the major objectives.

Another important aspect of the studies that have been done concerns lower-level factors that might be exhibited among the items of the Gumpgookies test. The hope was that such a series of analyses might indicate components of motivation that were perhaps under-represented or over-represented in the total score. The reader has already been alerted to the fact that young children may be affected, in part, by the position of the right answer to an item, by the position of an item in a test, or

by the order in which an alternative for an item is expressed. The total score on current forms of the test, it is stressed, should not be affected systematically by such influences, for reasons reiterated below.

Items for the three forms of Gumpgookies--a 75-item, individually-administered form; a 100-item, group-administered form for non-readers; and a 100-item, group-administered form for readers--were assembled in such a way that approximately equal numbers of items had the correct answers in the left, right, up, and down frames of the illustrations. Some illustrations were also placed in the up-left, down-left, up-right, and down-right positions. The instructions to examiners also specified the order of describing illustrations for the individual and the group form for non-readers, in such a way that the possible influence on the total score of a tendency to choose the first answer described or a tendency to choose the second answer described would be minimized. For a small number of items in both forms (five), the separate illustrations are not described, as in items that pose such questions as "Which is your gumpgookie climbing the tree?"

Because of various circumstances that will not be explicated, the desired equality with respect to all possible comparisons that might be examined was not achieved exactly. Nevertheless, the numbers of correct answers in the main answer positions--right, left, up, and down--are very nearly the same, as are the numbers of items that could reflect either primacy or recency.

With the form of test used for readers, as in the fourth-grade sample, no control could be exercised over the possible influence of primacy versus recency, i.e., over which illustration was described first. The supposition has been that the descriptions appearing at the left, up, up-left, or up-right positions would be read by the subject before those appearing at the right, down, down-left or down-right positions.

These precautions having been taken in designing the format of the test, the expectation would be that the total score would not be affected in any systematic way by tendencies of particular children to be influenced by answer-position preferences or by order of presentation of alternatives.

This chapter, following its lengthy preamble, will present some of the many factor analyses that have been made of the three randomized forms of the Gumpgookies test. The reader will be spared some ventures into cul de sacs as well as the details of test runs periodically made to assure that the programs and equipment were yielding correct results.

The analyses have veered back and forth between oblique and orthogonal rotations. Nevertheless, though arguments in favor of relatively independent factors at times seem persuasive, the latest inclination of the authors is towards oblique solutions, which, after all, can yield orthogonal factors if the data so dictate.

Factor Analyses for the Three Grade Groups

The first two group forms of Gumpgookies were administered to 212 first-graders, 223 second-graders, and 233 fourth-graders in Honolulu schools. For each grade group separately, a principal-axes solution for six factors was obtained and rotated to an oblique simple structure by a biquartimin solution with gamma equal to .5. Table 6.1 presents the three sets of rotated (oblimin oblique) factor loadings.

An Initial Approach to Super-Factors for the Three Grade Groups

Comparison of the factors across grades by inspection alone was almost impossible. Hence, at the suggestion of Paul Horst, the Pearson correlation coefficients among the 18 sets of factor loadings (six sets for each grade) were obtained. These are given in Table 6.2.

TABLE 6.1

Obliquely Rotated Factor Loadings for the 100-Item Group Form
of Gumpgookies for Three Grade Groups (First, Second, and Fourth)

Item Number	Factor Loadings*																	
	Factors for First Grade (N = 212)						Factors for Second Grade (N = 223)						Factors for Fourth Grade (N = 233)					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	12	45	08	22	06	-18	35	15	24	18	14	02	13	04	-09	45	-07	12
2	05	17	01	15	37	06	16	-01	-05	-12	02	42	19	08	-07	43	02	47
3	-13	24	10	19	07	-03	-07	03	00	32	-05	-05	00	-08	-07	-09	09	-01
4	-05	30	-07	27	24	06	-05	11	35	07	-05	15	-09	57	11	01	20	-09
5	30	32	-17	38	-08	28	-03	54	18	-21	-09	06	08	08	55	14	06	01
6	05	22	35	26	20	10	38	25	40	03	05	21	54	20	08	14	02	21
7	-02	48	07	19	30	-12	12	25	25	19	14	52	15	01	05	12	00	35
8	01	-11	-15	-18	-18	14	-23	03	09	15	-26	18	-17	21	14	-07	18	-04
9	18	10	-25	05	10	01	-20	05	16	03	-24	15	00	22	11	03	22	-19
10	06	39	11	01	14	-24	-07	34	18	24	-02	05	17	07	07	67	-17	44
11	38	41	-12	21	-01	45	01	63	16	-08	30	22	-02	16	71	01	16	-03
12	14	02	-06	-14	31	12	02	22	33	14	10	34	03	67	24	14	05	07
13	05	28	-08	31	-19	03	21	19	35	-14	44	12	33	09	-07	23	-12	23
14	30	61	-13	21	03	17	10	50	19	45	24	26	-01	23	09	40	-01	-19
15	31	54	-15	13	29	19	-06	43	17	19	-02	19	-03	14	28	-06	04	05

*Decimal points have been omitted.

TABLE 6.1 (cont.)

Factor Loadings*

Item Number	Factors for First Grade (N = 212)						Factors for Second Grade (N = 223)						Factors for Fourth Grade (N = 233)					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
16	24	53	07	31	-03	21	12	71	27	07	25	10	19	27	77	09	17	18
17	16	19	09	40	03	21	50	12	16	10	03	14	57	01	21	10	-01	10
18	45	39	-03	29	07	51	-06	54	-09	-12	-10	08	09	36	52	26	12	30
19	06	00	06	-02	18	-04	09	25	12	-15	-14	29	-19	08	23	-02	22	-09
20	19	37	40	34	-08	-18	62	14	16	-02	14	-05	51	-08	06	04	09	23
21	71	55	11	30	16	-11	02	34	30	14	48	09	01	-01	32	58	-12	24
22	33	22	04	17	-03	13	17	27	27	-14	-08	12	09	-19	19	-05	-06	22
23	48	50	06	25	-02	45	19	53	21	28	14	-05	24	20	43	20	-10	42
24	17	22	-15	01	10	18	06	28	13	-11	19	19	28	00	26	-17	-18	10
25	23	35	03	06	05	-06	28	33	19	31	05	18	17	24	01	-38	-10	02
26	29	70	17	30	-04	11	12	31	18	-02	44	23	16	02	35	31	-05	07
27	07	45	60	26	11	-01	65	15	35	09	06	17	71	05	14	12	03	24
28	08	32	05	12	17	16	-02	02	16	09	-23	08	02	22	18	28	18	-14
29	25	48	28	23	-10	-04	41	27	60	05	16	07	33	13	06	23	-02	40
30	18	33	-15	10	-05	-16	11	-08	-06	05	06	56	-22	24	-15	02	02	-03
31	33	83	09	29	14	04	01	26	34	37	-15	29	08	72	18	41	-01	-03
32	24	59	-02	41	-35	01	16	44	25	05	17	13	26	28	38	18	03	12
33	50	47	-05	16	03	49	08	64	30	-03	-09	15	08	01	71	-01	11	-02
34	22	07	05	-18	-16	-34	23	-15	-11	13	09	08	43	12	-16	12	-08	-16
35	21	35	49	20	03	-04	45	16	26	16	09	-07	67	08	06	11	09	20

*Decimal points have been omitted.

TABLE 6.1 (cont.)

Factor Loadings*

Item Number	Factors for First Grade (N = 212)						Factors for Second Grade (N = 223)						Factors for Fourth Grade (N = 233)					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
36	21	17	-25	-01	02	09	-25	14	12	08	-02	18	37	02	08	11	35	32
37	20	63	15	43	15	21	15	40	15	20	13	-03	17	-03	32	-06	16	34
38	31	07	14	02	-29	07	12	13	30	-02	00	02	21	-13	18	-05	01	24
39	32	47	-02	12	-10	01	07	04	34	-05	14	-09	11	06	-13	05	-20	04
40	39	47	14	04	-09	30	31	-07	18	00	-06	-04	31	21	12	33	-17	-01
41	14	02	01	-02	-36	01	23	14	10	09	11	01	21	02	13	12	-23	13
42	24	10	18	-25	-11	13	14	46	30	-13	-06	64	-03	04	29	25	-12	35
43	21	15	52	17	-28	-20	53	09	-02	-13	-02	-07	47	01	06	-08	-09	29
44	07	15	05	11	-07	-14	49	-01	05	-08	24	05	48	12	-03	22	-24	03
45	48	53	-02	17	-09	12	15	33	32	36	12	11	29	27	04	31	-16	01
46	51	55	-08	21	34	15	04	23	37	-38	13	23	09	00	02	-05	-10	03
47	61	60	08	24	28	-01	03	16	19	38	-08	23	04	53	-06	03	-04	33
48	27	14	34	-03	03	24	62	06	17	02	06	13	68	-09	06	07	07	22
49	08	-05	02	-15	-05	30	-04	30	08	-01	08	-09	-05	17	23	07	19	-25
50	04	27	-04	12	41	07	04	13	18	-13	27	06	07	15	06	-01	46	-03
51	11	21	44	24	11	16	68	14	03	-08	06	03	69	-07	07	00	-03	11
52	35	23	28	29	-05	-09	49	36	27	-16	29	05	28	10	-02	19	-01	17
53	08	23	18	51	08	-06	45	12	22	01	17	-04	54	12	14	03	-11	-01
54	06	13	04	21	21	-04	19	17	22	17	-08	14	-30	00	-13	13	03	07
55	00	-09	39	-20	-08	-09	27	-15	-35	-10	04	03	30	-06	-03	05	-16	16

*Decimal points have been omitted.

TABLE 6.1 (cont.)

Factor Loadings*

Item Number	Factors for First Grade (N = 212)						Factors for Second Grade (N = 223)						Factors for Fourth Grade (N = 233)					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
56	36	40	00	07	40	11	05	-09	04	-04	02	43	-11	40	-05	17	46	-03
57	24	42	39	40	12	04	41	07	11	00	22	11	61	-05	17	04	-07	08
58	12	23	-03	37	02	-03	20	06	-01	20	19	10	-04	03	-05	-04	09	03
59	36	55	-08	42	12	-27	19	24	30	-07	65	28	-01	31	09	-08	-03	10
60	61	47	24	21	17	06	34	20	51	03	18	14	19	42	30	16	07	-02
61	32	22	17	18	-11	14	10	04	16	-11	20	05	01	-15	22	-04	19	00
62	26	34	-16	18	32	16	04	43	52	07	15	30	24	13	25	-06	30	15
63	50	40	34	35	29	09	65	22	14	-06	08	20	68	-09	16	09	-04	29
64	-04	-06	-51	-05	-07	15	-41	11	23	-06	-06	17	-11	51	07	02	30	03
65	-08	-09	21	-08	-02	-04	07	22	18	21	-01	05	18	05	25	02	-01	44
66	40	29	-13	29	34	25	27	42	19	05	27	30	24	04	02	28	03	29
67	11	08	-21	19	37	20	-15	12	11	15	-03	31	-05	29	29	26	36	12
68	66	61	05	17	-01	07	12	33	39	-06	03	14	21	40	17	30	31	25
69	53	27	23	27	24	-19	42	28	18	-14	42	16	38	00	02	26	-05	31
70	22	22	03	51	18	25	22	42	23	13	06	06	22	35	15	04	26	-09
71	58	37	18	33	09	15	22	23	57	02	19	10	12	-01	07	21	-05	08
72	10	28	04	10	-01	-05	-01	01	30	29	11	06	-03	51	12	11	16	-20
73	58	19	07	11	-05	03	04	18	27	43	14	44	06	00	-06	42	03	-08
74	53	46	-03	20	14	22	-11	32	52	-08	-17	34	05	26	36	41	44	-22
75	36	21	-05	29	-12	10	31	61	18	-07	22	14	20	25	37	37	42	05

*Decimal points have been omitted.

TABLE 6.1 (cont.)

Factor Loadings*

Item Number	Factors for First Grade (N = 212)						Factors for Second Grade (N = 223)						Factors for Fourth Grade (N = 233)					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
76	34	25	16	06	-31	-28	19	-13	-10	12	11	-31	14	-13	03	02	-33	03
77	18	15	02	51	00	-17	07	05	12	26	55	12	-02	00	26	55	-07	08
78	22	32	45	31	18	-44	58	13	30	-06	15	25	58	-03	-09	22	-18	06
79	55	52	05	26	13	30	11	69	37	19	09	12	16	27	33	00	41	06
80	50	58	17	46	-05	07	36	24	27	-35	06	22	16	02	25	33	24	02
81	-02	-07	60	15	17	-10	63	-01	10	06	14	10	36	00	06	03	-25	-01
82	69	55	-02	14	02	12	18	52	32	09	-16	50	14	00	22	21	30	40
83	37	38	-00	23	-15	02	12	22	16	11	07	-16	-07	29	03	-05	09	-04
84	23	11	03	06	-07	-34	36	14	08	-27	-04	10	52	12	-08	23	-12	09
85	01	10	-23	12	06	-14	-08	32	23	22	-24	13	-15	26	14	05	25	19
86	60	38	29	23	11	-04	29	16	41	-03	20	-03	30	-12	08	11	01	03
87	04	29	-11	24	18	-01	-18	00	11	12	19	28	05	44	18	35	56	-08
88	56	39	-16	45	06	-16	-03	04	13	20	34	01	02	20	15	00	12	-14
89	08	12	20	36	02	-13	35	03	21	13	10	-13	30	-13	10	-02	02	13
90	20	22	42	17	-20	-11	61	12	00	-08	-08	10	44	-05	05	05	07	12
91	48	23	-03	26	-07	-02	24	31	51	10	22	07	11	04	38	09	52	03
92	28	37	03	52	-13	17	02	45	44	-04	17	43	17	16	64	35	17	-07
93	56	36	-14	31	-21	04	18	32	50	-11	-17	18	28	41	-06	42	06	-05
94	57	30	-21	21	-07	02	22	43	19	-44	13	25	01	-01	00	44	09	09
95	24	21	15	39	-08	-09	30	16	49	03	17	10	14	21	02	-03	19	09

*Decimal points have been omitted.

TABLE 6.1 (cont.)

Factor Loadings*

Item Number	Factors for First Grade (N = 212)						Factors for Second Grade (N = 223)						Factors for Fourth Grade (N = 233)					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
96	20	20	-15	31	-19	-29	-02	16	13	-06	37	-02	06	27	20	36	15	33
97	34	21	15	32	-07	-17	05	11	37	23	13	11	10	27	-09	02	05	37
98	51	23	-18	09	-02	20	02	14	61	18	08	18	-01	51	11	03	24	02
99	41	03	-10	-13	-05	16	24	02	08	30	34	-04	26	43	01	28	-08	12
100	58	17	13	17	16	05	02	27	43	08	16	38	35	13	-06	39	-20	03

*Decimal points have been omitted.

TABLE 6.2

Correlations Among the 18 Sets of Oblique Factor Loadings for the Six Factors for Each of Three Grade Groups (First, Second, and Fourth), Based on a 100-Item Group Form of Gumpgookies

Correlations* for																	
First-Grade Factors						Second-Grade Factors						Fourth-Grade Factors					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1st Grade	2nd Grade	3rd Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade	9th Grade	10th Grade	11th Grade	12th Grade	13th Grade	14th Grade	15th Grade	16th Grade	17th Grade	18th Grade
1	51	02	23	-02	23	02	39	40	-09	17	11	-00	06	17	25	04	01
2	51	02	51	17	17	02	42	35	09	17	12	01	15	24	23	02	06
3	02	02	13	-03	-27	78	-20	-06	-06	08	-24	66	-44	-15	-07	-38	26
4	13	13	11	-03	-03	24	27	28	-04	38	-08	16	-05	20	09	09	02
5	-03	-03	11	15	15	-10	04	09	09	-01	38	-08	15	-05	02	22	02
6	17	-27	-03	15	15	-26	51	15	-04	-22	14	-14	16	59	-08	31	-09
7	02	78	24	-10	-26	-12	-05	-22	21	-22	-22	76	-43	-23	-05	-41	23
8	42	-20	27	04	51	-12	39	-06	04	04	21	-10	08	66	16	21	17
9	35	-06	28	09	15	-05	39	04	08	08	20	-03	23	16	13	19	-01
10	09	-06	-04	09	-04	-22	-06	04	-01	-06	-06	-17	29	-10	09	-03	-03
11	17	08	38	-01	-22	21	04	08	-01	-11	-11	13	-12	02	14	-20	07
12	12	-24	-08	38	14	-22	21	20	-06	-11	-11	-23	19	05	26	22	03
13	00	01	66	-08	-14	76	-10	-03	-17	13	-23	-36	-36	-13	01	-36	30
14	06	15	-44	15	16	-43	08	23	29	-12	19	-36	08	08	15	34	-27
15	24	-15	20	-05	59	-23	66	16	-10	02	05	-13	08	-01	-01	32	-02
16	23	-07	09	-08	-09	-05	16	13	09	14	26	01	15	-01	-09	08	08
17	02	-38	09	02	02	-36	15	19	-03	-20	22	-36	34	32	-09	22	22
18	06	26	02	02	-09	30	30	-01	-03	07	03	30	-27	-02	08	-22	-22

*Decimal points have been omitted.

Inspection of the relations among the factors for the three grade groups revealed two factors that are clearly similar for the three groups. Table 6.3 highlights these relations.

As an aid in interpreting these two factors that were determined by inspection to cut across all three grade groups, a list was made of the factor loadings for each of the grades for items with the highest loadings (a sum equal to at least .50).^{*} Inspection of the items with the highest loadings for these two "super-factors" led to quite meaningful interpretations. The first factor, A, is loaded heavily with items that stress excellence and winning--for example, items embodying such contrasts as "doesn't know" versus "knows," "knows some" versus "knows more," "racing" versus "winning," "one of the class" versus "best of the class," "has some friends" versus "has more friends," and "just like others" versus "better than others." The emphasis on aspiration toward excellence is compelling, and the factor seems to involve primarily self-confidence in ability to excel. Aspiration toward excellence is accompanied by confidence in one's ability. The factor seems to be tapping an important component of achievement motivation, a conceptual component, whereby the individual sees himself as a high achiever.

Super-factor B, by means of a similar inspection of items with high loadings for the three groups, was at first interpreted as positive orientation to school. The best items reveal such contrasts as "likes school" versus "doesn't," "school is fun" versus "gets tired in school,"

^{*}This list is not included here, because in the meantime the test has been revised to reduce the number of items to 75. For some of the retained items, revisions of wording and of illustrations have been made.

TABLE 6.3

Two Sets of Correlated Factor
Loadings Across Grade Groups

		Grade Factor	1 <u>3</u>	2 <u>7</u>	4 <u>13</u>
Grade	Factor				
1	<u>3</u>		--	.78	.66
2	<u>7</u>		.78	--	.76
4	<u>13</u>		.66	.76	--
		Grade Factor	1 <u>6</u>	2 <u>8</u>	4 <u>15</u>
Grade	Factor				
1	<u>6</u>		--	.51	.59
2	<u>8</u>		.51	--	.66
4	<u>15</u>		.59	.66	--

"wants to stay in school" versus "wants to go home," "likes to learn" versus "likes to play," "helps before school" versus "plays before school," "stays to listen to the teacher read" versus "goes out to play," and "trying to write" versus "watching." The loadings of certain items, however, suggest that work enjoyment is a more apt description of this affective component of motivation, because some of the items refer to work in any situation.

Once the correlations for the factors involved in super-factors A and B had been extracted from the intercorrelations in Table 6.2, the next highest inter-grade coefficient was .40, between factor 1 in the first grade and factor 9 in the second grade. Neither of these had a correlation above .25 for any factor in the fourth grade. Although the interpretation did not seem so clear as for the first two super-factors, a tentative hypothesis was that this super-factor, C, represented interest in activity--in doing something, learning something, making something, telling, practicing, jumping, thinking, counting, etc.

Factoring the Factor Loadings To Get Super-Factors for the Three Grade Groups

In order to facilitate further comparisons among the factors for the three grade groups, the Pearson correlation coefficients among the factor loadings for the original, obliquely rotated factors (Table 6.1) were themselves factored. Two orthogonally rotated solutions were obtained, one with three factors and one with six. The latter was more fruitful, so that only it will be presented (see Table 6.4). In general, the three-factor solutions confirmed the three strongest super-factors in the six-factor solution.

TABLE 6.4

Orthogonally Rotated Factor Loadings Based Upon
 Pearson Correlation Coefficients of Obliquely Rotated
 Factor Loadings for First-, Second-, and Fourth-Graders
 (100-Item Form of Gumpgookies, N's 212, 223, and 233, Respectively)

			Super-Factor Loadings*					
			<u>A</u>	<u>B</u>	<u>C</u>	<u>D</u>	<u>E</u>	<u>F</u>
Initial Factors	First Grade	1	-02	40	44	08	-08	41
		2	05	36	28	29	12	58
		3	86	-14	-11	06	-03	07
		4	16	15	-11	01	10	83
		5	-01	-03	-10	12	87	10
		6	15	80	-11	04	11	-16
	Second Grade	7	86	-14	-06	-09	-07	21
		8	-04	81	26	-06	05	20
		9	-06	36	21	31	16	38
		10	-04	-13	02	80	-02	-04
		11	07	-22	14	-23	-14	70
		12	-20	10	41	-12	70	-11
	Fourth Grade	13	84	-04	-02	-02	-07	08
		14	-58	08	06	54	15	02
		15	-12	81	-10	-14	-08	11
		16	-08	-05	77	08	05	14
		17	-47	34	-34	01	38	06
		18	51	10	40	-16	11	-16

*Decimal points have been omitted.

The super-factors previously described, A--self-confidence in ability to excel--and B--work enjoyment, were clearly identifiable in this solution, A having its highest loadings on the initial factors 3, 7, and 13 and B its highest loadings on 6, 8, and 15. Two additional super-factors are revealed by the analysis: C, comprised of initial factors 1, 9, and 16; and D, comprised of initial factors 2, 10, and 14.

For confirmation of the original interpretations of super-factors A and B and for arriving at interpretations of super-factors C and D, an attempt was made to assign each test item to a single super-factor on the basis of its loadings on the initial factors comprising the super-factors. Nineteen items went with A, 11 with B, 21 with C, and 14 with D. Only seven items could be associated uniquely with super-factor F, comprised of factors 4 and 11 but with no high loading among the fourth-grade factors.

Although nine items could be assigned to super-factor E, weighted with factors 5, 12, and 17, a fully satisfactory interpretation of it was not achieved. Several of the items, alone or in combination with others, suggested the ability to maintain attention in order to satisfy curiosity, but this interpretation was not highly compelling and was not confirmed by subsequent analyses based upon what seemed to be superior methodology.

Nineteen items were not strongly associated with any of the super-factors.

Items for super-factor C again suggest activity in a variety of settings. The activities involved transcend the school environment, although a small proportion are school-related or could occur in school. The content of several of the items brings to mind the idea of persistence. Overall, the factor suggests knowing what action to take, knowing the instrumental steps to an end. It is regarded as a cognitive component

of motivation, i.e., knowing instrumental steps. It will be referred to as instrumental activity.

Super-factor D, consisting of initial factors 2, 10, and 14, contains 14 items. Its interpretation is somewhat more difficult than for the other super-factors, perhaps partly because factor 2 is more highly loaded on super-factor F (for which factor 4, however, has a much higher loading and for which no interpretation has suggested itself). Observe, also, that factor 14 is negatively loaded on super-factor A, interpreted as self-confidence in ability to excel.

Almost all of the items assigned to super-factor D contain a reference to some model. Either the teacher or parents are present in the setting, i.e., watching the teacher, showing something to the mother, the teacher knows about him, telling his family, parents are sad. Some items also emphasize one's own independent accomplishments. The factor seems to involve ability to establish a goal, awareness of future goals, willingness to postpone gratification. It is tentatively interpreted as representing a purposive component of motivation to achieve, responsiveness to future goals. In some of the early factorizations of data for four-year-olds who had taken early forms of the individual test, a similar factor had suggested itself.

Other considerations have already led to reduction in the number of items and to revised wording in many of the items in the forms of the test used in the cross-cultural comparisons for four-year-olds and for other current purposes. Nevertheless, a further pursuit of unambiguous interpretations of super-factors for these first-, second-, and fourth-graders was regarded as potentially fruitful. It was decided to explore further the use of an orthogonal instead of an oblique transformation for the

original principal axes solution for each of the three grade groups. Because the test has been revised and because facilities for analysis of data were not limitless, however, additional work to interpret the resulting super-factors has not been undertaken.

Despite some further complications with reference to position influences, (i.e., reflecting a tendency to select answers in a particular location), primacy versus recency influences (i.e., reflecting a tendency to select answers heard first or last), and effects of item locations in the test, which are described later in this chapter, the first discovery of four identifiable super-factors, with the possibility of a fifth, was regarded as very encouraging. Although a few of the items (e.g., "needs help in getting dressed") are inappropriate for fourth-graders, as was anticipated, eventual production of a form that has factorial validity for children through age nine or possibly even 10 seemed assured. Hence one further type of analysis of these data for the three grade groups will be described later in this chapter.

Factor Analyses for the Ten Ethnic-Cultural Groups of Preschool Children

The 75-item form of Gumpgookies, with every feature possible randomized, was administered in the spring of 1969 to children averaging slightly younger than five years of age in 10 subgroups: 225 Mormons, 200 Catholics, 197 Jews, 90 Puerto Ricans, 202 Negroes (urban), 36 Whites (rural), 47 Hawaiians, 39 Orientals from the United States West Coast, 114 Mexican-Americans, and 234 American Indians. Additional data for Hawaiian Head Start classes which were tested in the fall of 1969 were later added for certain analyses to bring the total N for Hawaiians to 270.

The principal-axes method of factor analysis was applied to each sample, and the resulting 10 matrices of six factors each were rotated to orthogonal simple structure.

These matrices, not here reproduced, have yet to be fully studied in a number of respects. It is clear at this point that the factorial structure of the test is not identical for all groups, even those with the larger N's or those for which socioeconomic levels can be regarded as similar. Also, in the course of other aspects of the investigation, differences among groups in susceptibility to extraneous influences such as answer position and primacy versus recency in the presentation of item alternatives were suggested. Extensive work on the differences among factors for the subsamples must await further resources. In the meantime, similarities among the factors have been explored in various ways, and these are being reported.

Initial Approaches to Super-Factors for the Ten Ethnic-Cultural Groups

With the larger number of solutions to be condensed, the task by inspection alone was even more formidable than in the case of the three solutions for the 100-item form that had been given to first-, second-, and fourth-graders. A table was made, showing the items having the five highest loadings for each of the six factors for each of the 10 groups. Patterns indicating identical factors were almost impossible to discern by this method. Hence, resort was again had to a further application of factor analysis with a view to seeing whether or not there were super-factors cutting across the various ethnic-cultural groups.

The separate analyses had resulted in 10 six-factor orthogonal solutions based upon the phi coefficients among the 75 items. Orthogonal

rotation was selected with a view to possible reduction of problems inherent in correlated factors within each subsample. On the other hand, when this method later was tried for the three-grade solutions, after the analysis based upon correlations of obliquely rotated factors had been completed, the solution on the whole was less satisfactory.

Many negative factor loadings remained even after appropriate reflection of factors having a majority of negative loadings. This suggested that some of the super-factors that might be obtained by factoring the correlations among the initial factor loadings for the six factors for each of 10 groups might be most meaningful if bipolar super-factors were accepted. Indeed, this possibility was fully explored during many hours of arduous work. It was finally rejected--at least as far as this report is concerned--for two reasons. With the time and resources available, problems were insurmountable in trying to interpret bipolar super-factors.

A second reason for rejection of bipolar super-factors at this stage of the development of the test was that logically they would have implied different scoring keys for the several subsamples. One response on a particular item would have been credited with a point for some of the subsamples, whereas the alternative response would have been awarded a point for other subsamples. Perhaps this is the way it has to be, eventually. But one of the major purposes of this venture into trying to test the motivation of preschool children to achieve in school is to discover whether a single test can be serviceable. For this reason, combined with the formidable analytical problems, bipolar factors have been put aside for the present.

Awaiting output of computing equipment, the investigators used various inspectional and tallying devices in an effort to anticipate super-factors that might emerge. Although such visual perusal of the 10 six-factor solutions produced some hints as to the composition of two or three super-factors, no assurance of the validity of interpretations was possible with such a plethora of numbers to compare and contrast.

When an orthogonally rotated solution for six super-factors, based upon factoring the intercorrelations among the 10 sets of six factor loadings, was available, it, too, presented problems. Because ethnic-cultural differences had been expected, however, these difficulties were no cause for surprise. In a first approach to identification of super-factors based upon this factorization, an effort was made to locate for each of a possible six super-factors one of the initial six factors for each subsample that would have (a) the highest loading on that super-factor for the subsample and (b) coincidentally, the highest loading for that factor among all of the six initial factors. In this first approach, high negative loadings were accorded equal status with high positive loadings. One problem was that the initial factors did not behave in a completely orderly way. This is not surprising, because of general problems in factor analyses of binary data and the small numbers of cases for some subsamples. In addition, the complexities of interpreting "double-negatives," coupled with an aim to have one scoring key rather than 10 different ones, led to the seeking of a super-factor solution that depended only upon positive loadings on super-factors for the 60 initial factors. The anticipated outcome was that some super-factors would not be represented for some of the subsamples if only

highest positive loadings in both directions were admitted. This forecast proved to be correct.

For every super-factor, a highest positive loading for both dimensions was sought. If such a loading was not available or if instead a high negative loading asserted itself, note was taken. The listing of subsample factors to define each super-factor, however, was based only upon positive loadings. Hence some subsamples were not represented on some super-factors. And, for some super-factors, a particular subsample may have had its highest initial factor loading (i.e., for the row) on some other super-factor, even though the column loading for that factor was the highest among the six loadings in the column. In such a case, the subsample initial factor was not included for the super-factor. Any one initial subsample factor thus was regarded as contributing to only a single super-factor--this despite the fact that in several cases the question of assignment of initial factors to super-factors was debatable.

The separate six-factor solutions, each for one of the 10 subsamples, are not presented herein.* The first six-factor, super-factor solution is given, however, in Table 6.5; it is followed by a summary of the factor numbers and super-factor loadings for each subsample group in Table 6.6.

One strongly represented super-factor, identifiable for all 10 groups (super-factor I), was associated with 19 items. A second, less compelling, was also represented for all groups, and it had 14 items loaded on it (super-factor II). A third, III, which was associated

*Had this course been chosen, our able secretaries might be with us no longer.

TABLE 6.5

Orthogonally Rotated Factor Loadings on Six "Super-Factors"
of the Six Initial Orthogonally Rotated Factors for Each of
Ten Subsamples (Based on 75-item Gumpgookies)

		<u>Super-Factor Loadings</u>					
	Initial Factor Nos.	II (1)	I (2)	III (3)	V (4)	IV (5)	VI (6)
Mexican- American	1	-.75	-.14	.04	.28	-.06	-.07
	2	.12	.49	-.56	.20	.04	-.03
	3	-.64	.20	.12	-.32	.03	.03
	4	.02	-.36	.10	.02	-.22	-.13
	5	.36	.07	.66	-.04	-.09	.38
	6	.47	-.13	-.30	.06	.14	-.22
Negro (Urban)	7	.65	-.08	.22	.04	-.05	.31
	8	-.09	.53	.11	.22	.26	.06
	9	-.36	-.39	.07	-.46	.07	.12
	10	.04	.02	-.41	.40	.02	-.03
	11	-.24	.13	-.33	.09	-.03	.29
	12	.15	.32	.18	-.06	-.13	-.09
White (Rural)	13	.40	.38	.12	-.18	.13	.15
	14	-.45	-.07	.22	-.19	.36	.09
	15	.08	-.03	.15	.46	.31	-.15
	16	.09	.59	.04	.02	.07	.08
	17	-.14	.10	-.09	.04	-.16	.50
	18	-.17	.05	.34	.15	.21	.19
Oriental (West Coast)	19	-.01	-.02	-.09	.12	-.37	.43
	20	-.38	-.09	-.03	.03	.10	-.27
	21	.41	.46	-.03	-.25	.05	.01
	22	.02	-.14	.08	.01	.13	.31
	23	.18	-.47	.13	-.09	-.02	-.05
	24	.28	-.04	-.41	.03	.14	.18
American Indian	25	.45	.05	.42	-.06	-.25	.30
	26	-.66	.06	-.00	-.03	-.01	.11
	27	.13	.10	-.35	.29	.53	.04
	28	.06	.37	-.30	-.25	-.12	.12
	29	.12	.47	-.12	-.10	.16	-.17
	30	.50	.08	-.00	.01	-.12	-.17

TABLE 6.5 (cont.)

		<u>Super-Factor Loadings</u>					
	Initial Factor Nos.	II (1)	I (2)	III (3)	V (4)	IV (5)	VI (6)
Mormon	31	-.10	.56	.26	-.06	-.15	.05
	32	-.33	-.10	.05	-.39	.35	.42
	33	-.12	-.24	.23	.16	-.24	.20
	34	-.00	.07	-.21	.53	.06	.06
	35	.69	.16	-.01	-.09	.14	.07
	36	-.01	-.08	.05	-.09	-.58	-.18
Catholic	37	.28	.52	.29	.15	-.32	-.06
	38	-.57	.12	-.07	-.17	.09	.00
	39	.31	.07	.02	-.50	.33	.24
	40	.46	.02	.18	.22	.20	-.14
	41	-.23	-.32	.02	.11	-.03	-.46
	42	.28	.14	-.05	.43	.07	.45
Jewish	43	.18	.10	.08	-.12	.07	.66
	44	.01	.60	-.41	.06	-.02	-.16
	45	-.02	-.25	-.08	.72	.02	.17
	46	-.61	-.02	.05	-.13	-.22	-.08
	47	-.12	.05	.26	-.17	.57	-.14
	48	.56	.23	.38	-.21	-.17	.13
Hawaiian	49	.70	-.04	.12	.19	-.25	.24
	50	-.08	.32	-.36	-.14	.13	.04
	51	-.14	.45	.23	.12	-.00	.30
	52	.09	-.18	.62	-.14	-.19	.10
	53	.38	.02	-.17	-.10	.56	-.24
	54	-.44	.28	.19	.31	.11	-.07
Puerto Rican	55	.32	.11	.55	-.16	.24	.15
	56	-.20	.47	-.33	-.07	.10	.09
	57	.24	-.13	.20	-.22	-.55	-.00
	58	.34	.06	.14	.46	-.12	.01
	59	.06	-.09	-.53	-.07	-.09	.11
	60	-.45	.11	-.12	.06	-.09	-.43
<u>Eigenvalues</u>		7.79	5.07	4.16	3.17	2.73	2.52

TABLE 6.6

Initial Factor Numbers for the Positively Loaded Factors
Associated With Each Super-Factor for the Ten Subsamples,
Together With the Loadings on the Super-Factors

	I		II		III		IV		V		VI	
Mexican-American	2	.49	6	.47	5	.66						
Negro (Urban)	8	.53	7	.65					10	.40	11	.29
White (Rural)	16	.59	13	.40	18	.34	14	.36	15	.46	17	.50
Oriental (W. Coast)	21	.46	21	.41							19	.43
American Indian	29	.47	25	.45	25	.42	27	.53				
Mormon	31	.56	35	.69	31	.26			34	.53	32	.42
Catholic	37	.52	40	.46			39	.33			42	.45
Jewish	44	.60	48	.56			47	.57	45	.72	43	.66
Hawaiian	51	.45	49	.70	52	.62	53	.56				
Puerto Rican	56	.47	57	.24	55	.55			58	.46		

with 14 items, was present for six groups, excluding urban Negroes, Catholics, Jews, and West Coast Orientals. Super-factor IV, associated with 10 items, was based on positively loaded factors for five groups, excluding urban Negroes, Mexican-Americans, Puerto Ricans, Mormons, and West Coast Orientals.

The remaining two super-factors, V and VI, could be associated with too few items to permit their identification. Hence, especially in view of an improved solution that became available later, their interpretations were not pursued.

Even with the three fairly strong super-factors revealed by this approach, conclusions are provisional, attempting to take into account different representations of the various subsamples as well as contrasts with the super-factors obtained for older children in school grades 1, 2, and 4 in Hawaii only, as described in the following section.

That the composition of the initial motivational factors for different ethnic-cultural-economic groups should differ is no cause for surprise or disappointment. Nor could one expect that life in complex and different environments would be so simple that factors comprising motivation to achieve in school would remain the same for children of different age groups. Especially would there be small likelihood that the factorial composition of the test would be the same for children of preschool age as for children in regular school attendance in grades through the fourth. Presumably, or perhaps even ideally, the interpretations of factors would be more general or undifferentiated for preschool children and could become more specific for older children.

It may be mentioned at this point that the initial intent of the very extensive factorial analyses of the test results for different

subsamples of the same age and for different age groups has not been to obtain scores on separate factors. Rather, the purpose has been to gain insight into the constituents of motivation for children of different backgrounds and different ages, and thereby to foster increased understanding and subsequent fruitful research on ways to increase motivation. Nevertheless, scores on separate factors comprising motivation to achieve appear now to represent an ultimately attainable goal.

Comparisons of Initial Super-Factors for the Three Grade Groups and the Ten Ethnic-Cultural Groups

Although overlapping of the initial super-factors for the 10 subsamples of younger preschool children (mean age 58 months) and those obtained for Hawaiian children in the first, second, and fourth grades is far from one-to-one, nevertheless there are resemblances. Table 6.7 shows the relation of the initial super-factors for the 10 groups of preschool children with those for the three elementary school groups. It also shows with which super-factor, if any, each of the 75 items was associated. The strongest four super-factors in each set have 38 items (out of the possible 75) in common. The comparisons will not be pursued in detail, because of more promising results of a different methodology, to be presented later in this chapter.

Super-factor I, the strongest, has four items in common with super-factor A, identified as self-confidence in ability to excel; four items in common with super-factor B, identified as work enjoyment; and five items in common with super-factor C, identified as instrumental activity. Super-factor II also has items in common with super-factor A (self-confidence) and C (instrumental activity). Super-factor III has appreciable positive loadings for six subsamples. It does not have

TABLE 6.7

Relation Between Two Initial Super Factor
Solutions, With Associated Numbers of Items

		<u>Initial Super-Factors for Grades 1, 2, 4</u>							No. items
		A	B	C	D	E	F	None	
Initial Super-Factors for 10 Groups of Preschoolers	I	4	4	5	1			5	19
	II	2		4	2	2	2	2	14
	III	2		5	2		3	2	14
	IV		5		2	2	1		10
	V			1		1		3	5
	VI		1		1	1		1	4
	None	2		2	2		1	2	9
No. items		10	10	17	10	6	7	15	75

appreciable positive loadings for the Negro (urban) nor the small Oriental (West Coast) sample, and its loadings for Mormons and Catholics are not high. Its items stress trying to do well, continuing with a task, trying to write, liking to tell stories versus listening, counting flowers, getting smarter versus tired when counting.

Super-factor IV appeared at first inspection to be concerned with positive orientation toward school. It has high loadings for six items associated with super-factor B, interpreted for the sample of first-, second-, and fourth-graders as work enjoyment. In the course of attempts to interpret super-factor IV in relation to super-factor B, it became apparent once more that at this level of analysis influences of answer position, item position in the test, or primacy versus recency as to order of presentation of alternatives, or combinations of such influences, were affecting the factors. Although such possibilities had been explored for each subsample factor analysis and no compelling evidence presented itself, when super-factor IV was compared with super-factor B and the items selected on the basis of all data to identify these factors were examined, it became clear that both super-factors tended to be loaded with items in which the keyed answer was presented first. For the 10 items in the 75-item form that were loaded on super-factor B, seven reflect influence of primacy. There were, however, 38 items that could have pulled together on one factor if this were the only influence. Likewise, for super-factor V, although eight of the 10 items identifying the factor reflected primacy, again 38 items could have shown this effect but did not.

Since after this analysis had been done it became apparent that factoring the correlations based on the 10 groups combined would be

more promising, no further attempts were made to interpret the content of these initial super-factors.

An Improved Approach to Super-Factors for Grades 1, 2, and 4

It had become apparent that a more clear-cut approach to super-factors might be to factor a matrix of phi coefficients computed for all relevant groups combined. Accordingly, a factor analysis based upon the combined samples of first-, second-, and fourth-graders in Hawaii was done. The rotation specified was orthogonal, with again six factors to be extracted.

This solution on the whole confirmed the super-factors previously discussed and yielded a structure that seemed more clearly interpretable than that previously available from the factorization of the correlations among factor loadings for the three grade groups.

As before, a super-factor, A', strongly identified as concerned with excellence and having 23 items associated with it was manifest. It can be interpreted as self-confidence in ability to achieve. The individual with a high score on this factor sees himself as a high achiever. The most heavily loaded items reveal no recency influence, though it shows up when the less strongly loaded items are added. This recency effect is perhaps interpretable as a tendency to withhold response until both gumpgookies for each item have been described. The most heavily loaded items, however, which emphasize being the best in the class, being better than others, being able to write just like teacher, knowing more than others, and winning a running race, do not reflect a recency effect. The items associated with this factor were scattered over the test, from item six to 95.

A second clear super-factor, B', which on the basis of the earlier analysis was thought possibly related to general work orientation, is more definitely associated with work enjoyment. It involves positive affect toward work. Among its 17 items, those that do not explicitly mention school or learning or studying nevertheless depict work, with few exceptions. This super-factor, like A', was scattered over all items (from item five to 92). Although among the highest six loadings five reflected primacy, only six of the top 10 and nine of the top 17 revealed this tendency. Hence it is concluded that this factor is not dominated by primacy or recency influences. Nor does it appear to be affected by position preferences.

A third strong super-factor, C', which is associated with 23 items, is interpreted as involving instrumental activity. The items loaded on it suggest initiative--being active or eagerly ready to engage in activity, energetic, not tired, able to do something, willing to try. Subsequent analyses, factor by factor and item by item, suggest that, although there is adequate reason to describe such a factor, it can reasonably be presumed to manifest itself through some different activity preferences for different age groups. Hence different items for the three groups are loaded on the factors that have loadings of .35 or above for the super-factor.

Two characteristics could not escape notice. The first is that the items loaded heavily on this factor tended to be those in which the keyed answer was read first. This suggests that the child motivated to be active and energetic tends to make a decision quickly--to choose the answer that he first hears described by the examiner.

Another striking finding was that, among the 23 items that could be uniquely assigned to this factor, 19 were in the last half of the test of 100 items and 12 (or over 50%) were in the last quarter. This was not the first time in this investigation that a factor apparently had been associated with position of items in the test. And the hypothesis of persistence to account for factors heavily loaded with items toward the end of the test had been discussed many times. In this instance, the hypothesis of instrumental activity had been deduced from the content of the items before their positions in the test had been examined. Nevertheless, the reality of the finding of higher loadings for items toward the end of the test is inescapable, and it is suggested that this tendency lends confirmation to the interpretation of the factor as persisting in initiating goal-directed activities.

Super-factor D', with 12 items, is not so strong as are A', B', and C'. As was found for super-factor D, its items suggest orientation to future goals, identification with a model, awareness of future implications, and delayed gratification. It is tentatively interpreted as responsiveness to future goals. It shows no answer-position effects but a tendency toward primacy effects. Eight of the 11 items ascribed to this factor are in the first half of the test. For a normal curve approximation to the binomial with $n = 11$ and $p = .50$, the z value corresponding to 7.5 is 1.21. Hence the finding is not inconsistent with the chance hypothesis.

A fifth super-factor is designated E'. It is not particularly associated with super-factor E, based upon factoring correlation coefficients among factor loadings for the separate analyses for the three grade groups. The major emphasis of the 13 items assigned to this factor is on quality of work versus amount--doing work right versus a lot, doing

well versus a lot, always doing one's best, getting problems right versus doing many, thinking versus watching, thinking versus waiting. This factor is about equally evident for the three different grade groups. Its items imply an internal standard of excellence--the ability to evaluate one's own work. It will be referred to as self-evaluation, an ethical component of motivation.

There is some evidence of a possible primacy influence, for all five of the top five items and nine of the total 13, and rather less evidence for a possible preference for the down versus up answer position.

Super-factor F' has loadings for too few items (six) to warrant firm interpretation. The items, though few in number, seem to emphasize self-reliance and dependability.

Six items were perhaps not legitimately assignable to any of the six super-factors, though they were placed with the super-factor on which they had their highest positive loading.

Comparisons of Factors for Grade 1, Grade 2, Grades 1 and 2 Combined, and Grade 4 with the Super-Factors for Grades 1, 2, and 4

In a somewhat different look at the super-factors for the elementary school children, items in the 100-item forms were first assigned to one of six super-factors based upon analysis of the three grade groups combined. Each item was assigned to the factor with the highest positive loading. For each factor, the items were listed in order of loading. Then separate recordings were made to show on which factor for four additional analyses each item had its highest loading: for first grade alone, second grade alone, first and second grades combined, and fourth grade alone.

For the items also included in the 75-item form used for the 10 ethnic-cultural groups, again recordings were made to show on which factor for each of the 10 separate analyses each item had its highest loading.

Super-factor A', interpreted as self-confidence in ability to achieve, was equally clear for grades 2 and 4 and almost as strong for grade 1. Its identification is in little doubt, with 21 items loaded .33 or higher. No one of the 10 groups of younger children had the items on this super-factor concentrated on a single factor. Rather, the items were scattered over several factors.

Super-factor B', work enjoyment, was most strongly identified for second-graders. For first-graders, the items for B' tended to be divided between two factors rather than falling on a single one. For the combined first- and second-graders, nine of the top 10 B' items were on a single factor, and nine of the 10 B' items were on a single factor for grade 4. It appears that first-grade children have not yet developed as strong a positive orientation toward work as second-graders. Perhaps fourth-graders already have tended to become disenchanted with school and hence with work. Comparisons of scores on this factor for the three groups might shed some light on this finding.

Super-factor C', related to activity, shows up most clearly for grades 1 and 2 combined. For grade 1, some of the C' items are more highly loaded on the factor related to work enjoyment. However, 19 of the top 20 items on C' could be regarded as having at least possibly significant loadings on a particular grade 1 factor, even though four had higher loadings on a second grade 1 factor. For grade 4, eight of the top 13 C' items were associated with a single factor. Some of the

C' items are not very appropriate for fourth-graders (i.e., needs help in dressing versus dresses itself, mad at itself because it can't write its name).

There is definite overlap though not complete identity between C' and a super-factor based on analyses for the 10 groups of younger children. When the factors were tallied with respect to the items of C', the groups most clearly having a single one of their respective factors associated with C' items were, in order: Jews; Hawaiians and Mexican-Americans (about equal); American Indians; Catholics; rural Whites, urban Negroes, and Mormons (about equal); Puerto Ricans.

Curiously, for a group of four items with moderately high loadings on C', three of the analyses for younger children show a shift from the factor associated with the still higher loadings on C'. These three groups are the Hawaiians, Mormons, and rural Whites. For these groups, the initial factors in question are associated primarily with super-factor II, whereas the initial factors that go with the more highly loaded C' items appeared mainly on super-factor III.

Factor D', a super-factor not so clear as A', B', and C', tentatively interpreted as responsiveness to future goals, appeared to be more definitely associated with a single factor for first-graders and for fourth-graders than for second-graders. However, it was primarily associated with a single factor for grades 1 and 2 combined.

Factor E', interpreted as self-evaluation, was perhaps most clearly identified for grade 2 among the three separate grade groups, though it is equally clear for grades 1 and 2 combined. For grade 4, the super-factor E' items seem to split mainly into two factors instead of staying together on a single one. This finding will require further study.

Super-factor F', with only six items and only two of these having loadings above .30, is not strongly enough identified to be interpreted with any confidence. Self-reliance or dependability were suggested. Nevertheless, five of the six items were together on a single factor for grade 1 and for grades 1 and 2 combined. Items on both D' and F' tended to be associated with the same factors for first-graders and for first- and second-graders combined. Perhaps this has some relation to the fact that for the first two grades the D' items reveal some tendency toward a primacy effect, the F' items some tendency toward a recency effect.

Comparisons of Five Super-Factors for Grades 1, 2, and 4 With Three Super-Factors for the Ten Ethnic-Cultural Groups

Because it had seemed possible to interpret five factors for the school-age children and fewer for the preschoolers, it was decided at one point to examine an obliquely rotated, five-factor solution for all of the older children and an obliquely rotated, three-factor solution for all of the preschoolers. There was then a possibility that second-order analyses might be enlightening, especially in relevance to the development of motivational factors.

Somewhat disappointingly, the factors yielded by this approach were not so clear-cut as those resulting from six-factor solutions. Moreover, possible answer-position effects appeared more prominently for the younger group, one factor having answers for the most part up or left, one up or right, and one down; and recency or primacy influences seemed evident for three of the five factors for the older group.

A double-entry table showing the congruence of items in the 15 cells associated with this approach again confirmed the troublesome role

particularly of answer-position effects but also of primacy-recency influences, though to a lesser extent.

Without presentation of the details involved in inspecting many numbers and much clerical work that transpired, the conclusion is presented that this resort to a smaller number of factors served only to intensify problems of interpretation of factors and their relations for the two groups.

A second-order analysis of the five super-factors for grades 1, 2, and 4. Before discussion of this particular phase of the analysis is abandoned, however, mention will be made of some interesting results of a second-order analysis of the correlations among the five oblique factors for the total school-age group. Factors 1 and 3 had loadings of $-.92$ and $.83$ on a bipolar second-order factor. Factor 1 may have been affected by recency, 3 by primacy. But factor 2 also was probably affected by primacy and has a loading of $-.03$ on the factor. Hence the contrast is not that of primacy versus recency.

If the interpretations of the factors--self-confidence and self-evaluation--have any validity, then it seems to make sense that they are at opposite poles. The person who is highly self-confident is not one given to self-evaluation. Note, also, that the first-order factors were negatively correlated, $-.60$.

For the second second-order factor, factors 2 and 4 were highly loaded-- $.80$ and $.74$, respectively. These were factors interpreted as related to activity, perhaps instrumental activity, and to work enjoyment. Factor 5, which in this five-factor solution had no clear interpretation, was loaded $.40$ on this second-order factor. Factor 5 had only a small number of items with high loadings. They are suggestive of concern for

adult approval of one's own achievements, although such an interpretation was by no means over-determined. The high loadings of factors 2 and 4 suggest constructively-directed energy, initiative, and perhaps compulsivity.

Comparison of Six Improved Super-Factors for Grades 1, 2, and 4 With Six Improved Super-Factors for the Ten Ethnic-Cultural Groups

The final comparisons of factors to be presented were based upon two orthogonally rotated, six-factor solutions--for all of the 668 Hawaiian children in grades 1, 2, and 4 combined and for all of the 1,607 children in the ethnic-cultural samples combined.* The former analysis has been discussed previously (pp. 103-106).

When the question of comparing factorial solutions based upon different groups (by age or ethnic-cultural backgrounds) was first considered, attention was directed by Paul Horst to a technique developed by Bruce Bloxom for rotating to maximally similar solutions for a number of groups (Bloxom, 1968). Since one of the purposes was to look at inter-group differences as well as likenesses, however, a solution that would maximize similarities appeared to be logically indefensible. Hence an alternative suggestion of Horst, to factor correlations among factor loadings for different groups, at first was adopted. The super-factors that resulted from this technique have been discussed earlier in this chapter. The approach still seems to have much to commend it when the purpose is to survey group similarities and dissimilarities, but the

*Under the influence of arguments on both sides, both orthogonal and oblique rotations have been used. Although the authors' bias is towards oblique solutions, for strong factors it does not seem to make a lot of difference which type of solution is used.

more recent comparisons of factorizations based upon combinations of subjects has seemed even more fruitful.

At this point, the rationale for putting the groups together in the way in which they were combined warrants repetition. Although children of different ethnic-cultural groups and children of different ages surely differ in motivational factors, some bases by which to compare the different groups are needed. Hence the approach of this series of analyses has emphasized attempts to identify factors by means of which subgroups eventually could be compared.

The authors are not unaware of other approaches to estimating the congruence of factors (e.g., Barlow & Burt, 1954; Wrigley & Neuhaus, 1955; Cattell, 1969). A prior concern, however, has been to define two sets of factors with which various solutions can be contrasted. Later, the results of the numerous factor analyses that have been made can be compared systematically with the defined super-factors. At this point, enough comparisons have been made to justify an assertion that the various subgroups do indeed differ in the factorial composition of the test, as was to be expected.

It will be recalled that the five most clearly identified super-factors for combined first-, second-, and fourth-graders were

- A', Self-confidence, a conceptual response;
- B', Work enjoyment, an affective response;
- C', Activity, possibly instrumental activity, a cognitive response, i.e., knowing what action to take;
- D', Responsiveness to future goals, a purposive response;
- E', Self-evaluation, an ethical response.

The possibility also exists that F' represents dependability or self-reliance, though it has only six items loaded on it and only two loadings of .30 or higher.

It may also be noted that A' is affected to some extent by recency and C' by primacy, as mentioned earlier. The other factors for the older children do not show such effects; nor, as a whole, do the factors show answer-position effects.

For the total group of 1,607 preschool children, the following factors are very tentatively identified:

1. General constructive activity
2. Self-evaluation
3. Optimistic self-confidence
5. Persistence
6. Work enjoyment

Factor 4, although it has 11 items loaded on it, is not yet clear. Further study is needed with respect to all of these factors for the younger children and to their relation to factors that can be identified for the separate ethnic-cultural groups.

These factors for preschoolers are still affected by irrelevant influences, although fairly clear interpretations on the basis of content seem to be possible. Factor 1 tends to have answers in the left and up positions, with primacy influences. Factor 2 tends to have answers in the down position, with nine of its 10 items in the first two-thirds of the test. Factor 3 items tend to have answers in the right and up positions. Factor 4 answers tend to be up, factor 5 answers down and, to a lesser extent, left, with more recency items than primacy and five of its six items in the last half of the test. Factor 6, one of the two most prominent factors (along with 1) tends to have answers that are down and to a lesser extent left, with nine of its 10 items in the first half of the test.

From the factorial findings for the randomized form as well as those for the original format of the test, the investigators are convinced that the effects of answer position, primacy versus recency, and item position

upon the factors are genuine and not to be solely attributed to chance. In some of the simpler cases, straightforward statistical tests seem appropriate and yield conclusions that the effects in question are statistically significant, with $p < .05$ or $p < .01$. In other more complicated cases, especially those seeming to involve a preference for one of two or three answer positions, where this preference is observed on an a posteriori basis, there is still some uncertainty as to how the statistical significance should be determined. One proposed method of computing a p value in such instances appears likely to be quite time-consuming, especially when the frequencies of the different answer positions in the population of such positions, i.e., in the scoring key, differ. Despite the fact that the computations that may be necessary for such cases have not been made, the observed phenomena are in some instances too prominent to be dismissed as attributable to chance. On the other hand, there is also persuasive evidence from divergent sources that the test and at least some of the factors are measuring sources of variance other than artifacts or extraneous influences.

Table 6.8 summarizes the relations among these two sets of factors, not by item numbers but by numbers of items, since the interpretations, especially for the younger age groups, are still tentative and since a copy of the test will not be available to every reader. It should be remembered that the interpretations of the factors for the older groups were based upon more items, since they took a 100-item form instead of the 75-item form that the preschool children took. An additional fact that has complicated interpretation is that the fourth-grade children read the items for themselves, so that the primacy-recency determination for each item had to be made separately for the first- and second-graders and the fourth-graders.

Relation of Orthogonal Six-Factor Solution for Grades 1, 2, and 4 (N = 668) and Orthogonal Six-Factor Solution for Ten Ethnic-Cultural Groups (N = 1,607)

Factors for 10 Ethnic-Cultural Groups

Hypotheses Concerning Answer-Position, Item-Position, and Primacy
Versus Recency Influences

Some general and quite tenuous hypotheses will be advanced as possible explanations for effects of answer positions, item positions, and primacy versus recency.

First, it should be noted that the question of right- versus left-handedness of the subjects in relation to preferred answer positions has not yet been explored in the course of these studies. It may be that right-handed children tend to take answers to the right and left-handed children answers to the left. Data on handedness are being collected for the children taking the Gumpgookies test in the fall of 1969 and thereafter, so that this question can be studied.

In these studies, the answer-position problem has been more strikingly manifest for the preschool groups than for children of school age. For the former group of 10 subsamples, super-factor 3, interpreted as optimistic self-confidence, was found to have a very large percentage of answers in the up or in the right position. This suggests that optimism goes with "upness," an explanation that at first thought seems too simple but that may nevertheless warrant further scrutiny. It is also possible that handedness was having an effect.

A tendency to take the left-hand answer, aside from its possible relation to handedness that remains to be studied, may reflect the western cultural pattern of reading--hence looking--from left to right, even for pre-readers. Many of them, even those who come from impoverished families, will have had this tendency implanted. Granted such a tendency, perhaps the child oriented to activity for the sake of activity, the one who wants to "get going," the one who prefers a rapid to a slow pace,

with winning and who identifies with excellence tries to exercise care to see that he does win or excel. Super-factor 3, also interpreted as involving self-confidence for the Head Start children, is almost equally divided between primacy and recency effects, but shows a preponderance of items with answers to the right. This tendency also could be associated with care if the child tends to look first at the left-hand picture.

On the other hand, super-factor C', interpreted as related to instrumental activity, reflects a possible primacy effect for the first- and second-graders, as does super-factor 1 for the combined Head Start groups. The items reflect activity and energy, not careful attention. Thus it may be that the child who seeks activity, who likes to be doing a variety of things, who prefers a rapid pace, tends to select the first alternative described to him so that he can the more quickly get on to the next one.

Aside from super-factor 1, no strong tendencies toward primacy or recency effects were evidenced for the super-factors based upon the 10 subgroups of preschoolers, although there is a possibility of a primacy influence for super-factor 6, work enjoyment. Such effects may be present for some of the subgroups and not others. This question can be explored more fully at a later time. In the meantime, the hypotheses developed above should be regarded as highly tentative.

Additional Work Needed

The separate factor analyses for the 13 subgroups and certain combinations of them warrant more intensive effort. Although the results were used extensively in the initial attempts to arrive at super-factors, work concentrated on interpreting the many factors was not systematically undertaken.

with winning and who identifies with excellence tries to exercise care to see that he does win or excel. Super-factor 3, also interpreted as involving self-confidence for the Head Start children, is almost equally divided between primacy and recency effects, but shows a preponderance of items with answers to the right. This tendency also could be associated with care if the child tends to look first at the left-hand picture.

On the other hand, super-factor C', interpreted as related to instrumental activity, reflects a possible primacy effect for the first- and second-graders, as does super-factor 1 for the combined Head Start groups. The items reflect activity and energy, not careful attention. Thus it may be that the child who seeks activity, who likes to be doing a variety of things, who prefers a rapid pace, tends to select the first alternative described to him so that he can the more quickly get on to the next one.

Aside from super-factor 1, no strong tendencies toward primacy or recency effects were evidenced for the super-factors based upon the 10 subgroups of preschoolers, although there is a possibility of a primacy influence for super-factor 6, work enjoyment. Such effects may be present for some of the subgroups and not others. This question can be explored more fully at a later time. In the meantime, the hypotheses developed above should be regarded as highly tentative.

Additional Work Needed

The separate factor analyses for the 13 subgroups and certain combinations of them warrant more intensive effort. Although the results were used extensively in the initial attempts to arrive at super-factors, work concentrated on interpreting the many factors was not systematically undertaken.

Further study of the characteristics of each item assigned to each of the relevant "improved" super-factors (A', B', C', D', E', F', 1, 2, 3, 4, 5, 6) should be made for each of the 13 subsamples (the three elementary school groups and the 10 ethnic-cultural groups). At this stage of a series of studies, effort has been concentrated on likenesses across groups rather than on the differences that are evident. These differences should constitute the focus of some further work.

It should also be of value to obtain approximate factor scores for the more clearly interpretable factors and to compare the various subgroups on the several factors. No attempts in this direction have been made as yet. Despite the "noise" associated with extraneous influences--which, incidentally, must affect other tests for young children as well as Gumpgookies--, the factors seem worthy of further pursuit. Chapter 8 will offer some suggestions as to additional types of analyses that will help to clarify the extent of the effects of unwanted influences as distinct from those of item content per se.

CHAPTER 7

EXPLORATIONS IN TEACHING PRESCHOOLERS MOTIVATION TO ACHIEVE

The theoretical orientation for this research defines motivation to achieve as a combination of learned responses. It further suggests that if these response constituents are learned, then they may be modified by manipulation of causal determinants in accordance with principles of learning. The implication of this suggestion is that motivation to achieve can be increased in the classroom by control of appropriate determining variables.

Operant Conditioning of Questioning

Early in the research reported herein, exploratory work began to identify classroom variables that could be manipulated to teach preschoolers motivation to achieve. A preliminary study attempted to apply operant procedures to condition questioning, an instrumental behavior that may facilitate individual initiative in learning. Conditioning took place in individual sessions in which a stimulus was presented and reinforcement was contingent upon the child's asking a question about it. The stimuli changed sequentially across sessions from concrete objects to abstract words. The reinforcement also changed sequentially across sessions from edible candy to abstract information, i.e., the answer to the question.

Seventeen four-year-old children were randomly divided into two groups. The first group (N = 12) underwent operant conditioning for questioning behavior with reinforcement contingent upon the child's improving his ability to ask well-formulated questions. The definition of what constituted a question varied among the children and with each

child throughout the 10 sessions. In some instances, the experimenter was required to engage in extensive prompting, including the use of model questions; in others, she merely presented the stimulus and manipulated the reinforcement. For each child, however, reinforcement was administered for successive approximations to the desired questioning response. The second group (N = 5) completed six individual sessions with the experimenter. The stimulus, questioning, was neither modeled nor reinforced. This group was used in order to determine the extent of the effect of familiarity with the experimenter upon the child's questioning behavior.

In the initial sessions, neither group of children constructed well-formulated questions for any stimulus presented. A further comparison of the two groups indicated that the questioning behavior seemed to increase in quantity and quality only when operant conditioning and modeling were employed, although personality variables were influential in determining the spontaneity and ease with which questions were asked.

A Curriculum To Teach Preschoolers Motivation To Achieve

The congruence of these and other similar results with the hypothesized model of motivation to achieve encouraged efforts to create a curricular design that would teach the response constituents of motivation to achieve. A preliminary curriculum of this nature was written and used in three Head Start classes during the school year 1968-69. Once each month, the staffs conducting the classes met for a discussion of the previous month's curriculum and of that for the ensuing month. This regular feedback from the teachers and from the classroom visits of the research staff was helpful in providing a basis for inclusion or

exclusion of various ideas. While activities suitable for a four-year-old were essential, they had to dovetail with the underlying aspects of motivation that were being sought.

The curriculum was presented to the teachers in two sections. The first concentrated on the use of ongoing activities and the second consisted of additional, specified daily activities. The ongoing activities are considered to be by far the more important of the two in that they are attempts to direct the covert learning that occurs in the teacher-child interaction. The role of the teacher in modeling the desired responses, eliciting some approximation to them, and reinforcing such approximations requires a high degree of awareness and comprehension. It was thought that, through this subtle use of characteristically human models, stimuli, and reinforcement, learning of the covert expectations, concepts, and standards would be facilitated.

The specific activities were exemplary ways in which the teacher-child interaction was to have been capitalized on in the classroom. In addition, these activities were so structured as to insure each child the opportunity every day to make one of the covert responses and to be reinforced for having done so.

Both types of activities were highly individualized. Teachers were instructed to alter each activity to fit their own personality as well as the level of development and personality of the child. The illustrative curricular materials that follow are simply examples of basic elements of the activities, omitting the unique reactions and contributions of each teacher as she dealt with each individual child.

At the time the design of the curriculum was undertaken, motivation to achieve was conceptualized as having six identifiable types of response

constituents: affective, conceptual, purposive, cognitive, ethical, and personal. The personal responses were thought of as attributable to separate personality characteristics that facilitate the expression of motivation to achieve. These characteristics have since been subsumed under the first five response categories.

Each unit of the curriculum was presented to the teachers in three sections: its rationale, a description of ongoing activities, and a number of suggestions for specific daily activities. Their treatment was sometimes varied, however, especially in the format of the curriculum itself. In the first two sections, the emphasis lay heavily upon the theory, with general activities being mentioned. The written material made available to the teachers served as a basis for discussion and also provided a document to which the teachers could refer. The units of the curriculum were modified and amplified in the course of the discussions.

Attempts were made to get the teacher to introspect and to analyze her reactions to children and to situations, lest there be unwitting obstacles to the rapport that was essential to the curriculum. The specific application, after the presentation of examples, was left to the teacher. In retrospect, the approach for these first two sections of the various units of curriculum appears to have been too individualized. The remaining section was more activity-oriented. The activities came under close scrutiny of the teachers. Their modifications and suggestions were in many instances adopted and their criticisms welcomed.

The following excerpts from the materials made available to the teachers will serve primarily to illustrate the types of ideas that were presented and discussed before attempts were made to apply them. For the first two units, an example of one of the specific activities that were suggested is shown.

Unit I

Affective Responses

Rationale

Expectation of positive affect underlies all motivation of behavior. If a child expects that achieving in school will make him happy, he will possess one of the basic components of motivation to achieve in school. The ultimate goal of this unit of work is to begin to teach the child that he will be happy in achieving in school. Such an expectation of positive affect can be developed only gradually through specific types of experiences.

Ongoing Activities

Teachers are concerned about their children. This concern is expressed through their active attention to the well-being and happiness of each child throughout each day. The communication of this concern to the children should be on a feeling level. Children have various needs and ways of expressing them. It is impossible to dictate how teachers can be aware of and meet these needs. If the teacher is truly concerned about her children, however, she will do everything she possibly can to try to meet their needs. The sincerity of her efforts to try is the most important aspect of her actions.

The positive affect which arises in a child when he feels that he is respected as an individual is invaluable. Teachers should treat each child with dignity and appreciation. Children are to be regarded as having potentialities that will not be reached without special efforts. If real respect exists for children, teachers will always honor their word to a child, speak to him at a challenging level, and prepare for the child as he would for an adult colleague.

During this unit, the emphasis is on a warm emotional climate. Teachers should do all they possibly can to eliminate unpleasant situations. All types of conflicts and stress should be minimal. Although undesirable incidents are to be avoided, certain limitations in the structure of acceptable behavior should not be sacrificed. Children find security when limitations are firm and consistent.

Illustration of a Specific Activity

Today, teachers should introduce the special person to each child. This person should give his whole and undivided attention to one individual child at a time. The activity engaged in with the child is not important; it could include anything and may be nothing more than being side by side.

Unit II

Conceptual Responses

Rationale

One of the prime components of motivation to achieve is the concept of the self as an achiever. The way the child sees himself in the learning situation appears to influence how well he will perform. If the child conceives of himself as an achiever, he appears to work until he does succeed in learning. If, on the other hand, he regards himself as a failure, he appears to stop trying and does indeed eventually fail.

The self-concept is a learned covert response, reached as a result of experiences and exposures. The school environment is new to the four-year-old child; the way he learns to see himself in this setting will greatly determine his motivation for learning throughout his academic career. In a sense, every moment may be crucial in the formation of

the child's concept of himself in school. He is constantly searching his environment for a reflection of himself. Thus, the activities throughout the preschool day are of critical importance.

Ongoing Activities

Teachers should treat each child as a unique human being: they should use each child's name, particularly in sentences that describe his achievements; individualize greetings and farewells; notice changes from day to day; show understanding and patience when a child spills or breaks something; differentiate between the child, as a person who is always acceptable, and his behavior, which may not always be; realize that when a child talks he believes he has something to say and needs full attention; and give explanations that the children are entitled to have so that they know what is going on.

Teachers should expect each child to behave differently: they should try to eliminate mass responses to questions; conduct quiet conversations with each child about things that concern him; allow each child to move at his own pace; respect his freedom to choose even when his choices differ from their own; reinforce individual accomplishments more than group accomplishments.

Teachers should honestly believe that each child is an achiever in learning: they should entertain no doubt that each child will be able to perform well; structure each task to a level so that each child can see himself achieving; help each child to attain realistic levels of aspiration; and at the same time, pace accomplishments so that they continually become a little more difficult.

Teachers should provide each child with opportunities to see himself as an achiever; allow each child frequently to take home with him some

tangible evidence that he has achieved; and provide opportunities for children to talk about their accomplishments.

Illustration of a Specific Activity

Introduce the "I Can Do" box. Each child should be given a box with his name on it. Explain that there are cards that fit the box. Each card has a picture on it of something that can be done. Each child may have any card when he shows that he can do what the card depicts. The teachers hold up the cards and give all the children an opportunity to perform one of the activities successfully.

Unit III

Purposive Responses

Rationale

The focus of this unit is to develop purposive responses, i.e., to learn in a school situation with conscious goal-directed behavior.

As a result of feeling safe and happy in his environment and knowing that he can achieve, the child is now ready to learn how to set up purposes towards which to work within the school setting. The concept of motivation itself implies a direction or goal for behavior. Motivation to achieve and to do well explicitly refers to behavior which has, as a component, the purpose or desire to achieve and to do well. Children who are so motivated at school do show their purposive behavior in their actions. Their behavior goes beyond merely holding to purposes set up for them either explicitly by the teacher or implicitly by the school situation, to the children's being able to come up with purposes independently of the structure. Their ability to form purposes from relatively unstructured situations is crucial if motivation is to exist.

Ongoing Activities

Teachers should concern themselves primarily with fostering the establishing of goals by the child himself, rather than his accomplishing goals set for him by the teacher. The stress should be on creating purposes for performing rather than on the success in performance.

Teachers should continually model the process of forming purposes by consciously trying to verbalize this process in front of the children even for the smallest of activities undertaken.

As each teacher models purposive behavior, she should reinforce every effort a child may make to set up a goal regardless of the goal itself. Teachers should capitalize on opportunities to reinforce the formation of goals by each child when they occur during the day. Remember, it is the process of forming a goal that should be reinforced and not the actual goal selected.

Teachers should constantly ask each child to express his goal in each activity in which he engages. The child may not be able to verbalize his goal but should continually be encouraged to do so.

The teacher should begin each day by asking the children to help plan an activity for the day. The process of suggesting goals by each child should be reinforced, and the teacher should verbalize as much as possible the need for goals in the child's activities. As the goals are formed and the process reinforced, some of the goals should be written or pictured on the board and made reference to as the activity is accomplished during the day.

Unit IV

Cognitive Responses

Rationale

If a child is to be motivated to do well in school, he must also know how he can do well in school, i.e., he must know the instrumental steps that will help him to obtain his goal. He must be able not only to differentiate between work and goals and to understand that work comes before the goal, but also to understand that the goal must be kept in mind before the work can be set in motion in an orderly way. Being able to differentiate behavior that will help him toward goals and behavior that will hinder him from obtaining them is also essential. Most important, however, is the child's realization that he is personally in control of whether or not he will eventually reach his goal.

Ongoing Activities

In order to teach these cognitive responses, teachers must continually emphasize, both behaviorally and verbally, the following three ideas:

1. It is necessary to work to get what is wanted;
2. Some ways of working are effective in getting what is wanted, other ways are not; and
3. Each person is more likely to get what he wants if he will work for it.

Unit V

Ethical Responses

Rationale

Inherent in the concept of motivation to achieve is the presence of a standard of performance that functions in the evaluation and

direction of behavior. Those individuals who appear to be motivated to achieve know how to evaluate their own performance and seem to be impelled to try to remove any discrepancy between their performance and their concept of how good their performance should be.

Ongoing Activities

Teachers should model and verbalize the evaluation process. They should continually try to get their children to consider their work critically. The approach must be positive in reinforcing the child for what he has accomplished and must also be effective in pointing out how the particular performance could be improved. Once this process is internalized, the child will become his own effective self-reinforcer and self-detractor with respect to behavior essential to motivation to achieve.

The final section of the curriculum was aimed at integrating the components of achievement-motivated behavior. The most successful activities tried in earlier sections were used, coordinating all the response constituents that were selectively reinforced previously. It must be understood that motivation to achieve has been separated into types of responses for the purposes of emphasis within the curriculum and that in reality they occur many times simultaneously and interact with each other.

Evaluation of the Curriculum

This curriculum to teach motivation to achieve to preschoolers was applied in an exploratory way to three Head Start classrooms in Hawaii. The ideas were tried out and then either accepted or, in most instances, modified or discarded. Evaluation was continuous and was specifically sought in monthly sessions between teachers and researchers who had spent many hours in the classrooms. This ongoing evaluation was immediately incorporated into the design of the materials, which were then given another trial.

In general, the teachers were highly favorable in their overall evaluation of the curriculum. They were not always certain of the implications of the ongoing activities but found the specific activities of great value. Their primary problem was to put into actual use the human stimuli and human reinforcement concepts. This difficulty experienced by the teachers in putting into operation in the classroom the ongoing activities related to motivation was a source of concern among the research staff. It was their impression that while the teachers did improve in so doing, the training provided was not sufficient to develop competence in this area, on which the success of the motivation curriculum depends. Future endeavors might well include a number of sessions in which staff and teachers expose themselves to a class of real children with a master teacher and engage in observation, directed discussions, practice teaching, and critical evaluation.

Inasmuch as the outcomes of this curriculum are so contingent upon the ongoing interaction between teacher and child, it seemed imperative that some additional evaluation of the teachers in the classroom be made. For this reason, the research staff made periodic observations of

every classroom. Analysis of these observations seemed to provide evidence that in this study, as in numerous others, the results were confounded by the teacher's personality, thus further verifying suspicions that the teacher-training techniques were not sufficiently effective to optimally modify the extent to which the ongoing activities took place.

Attempts to quantitatively scale teacher performance were also made. A scale that was developed by the staff appears to have potential but lacks empirical justification; no validity or reliability studies have as yet been made on it. Although the scale was used in rating these teachers, the lack of knowledge concerning its effectiveness makes the results questionable. Nevertheless, there was a general decrease in the ratings assigned to teachers over the entire year. Part of this decline might be attributable to the increased sophistication of the observers; i.e., as the research staff continued in the project, they became more and more aware of the subtle influences operating in the teacher-child interaction. Inasmuch as the written descriptions of the teacher's performances by the same staff seem to indicate that there was indeed progress in the ability of the teachers to perform in the specified manner, to interpret the findings of the scale at their face value would lead to an obvious contradiction. With the data at hand, it is almost impossible to evaluate the relative effectiveness of the different teachers in presenting the curriculum. Such evaluation is essential in further studies of this type of curriculum. Better ways of solving this problem are needed.

Pre-test and post-test data were gathered for three experimental and three comparison classes. Pre-tests given in late September consisted of Gumpgookies and the Stanford-Binet. The form of the Gumpgookies used

in the pre-testing was the 100-item, left-right form that was later modified in an attempt to reduce the effect of extraneous influences, primarily answer position and primacy versus recency.

The post-testing began in early April. By this time, however, the revised 75-item, randomized form of Gumpgookies had been completed, and it was also administered to three of the six classes.

An analysis of variance of the pre- and post-test Stanford-Binet scores showed a significant gain for both experimental and comparison groups. Although the experimental group was higher in both pre- and post-test ($p < .11$), the rates of gain were not significantly different for the two groups, as the data presented in Table 7.1 indicate.

An analysis of variance of the 100-item Gumpgookies scores for pre- and post-tests for both groups (Table 7.2) again showed a significant gain for both experimental and comparison groups. Although there were no significant differences at pre-test time between experimental and comparison groups, the rate of gain was greater for the experimental group as compared with the comparison subjects ($p < .11$).

An analysis of variance of the scores obtained with the improved 75-item, randomized form of the Gumpgookies administered as a post-test to three classes (Table 7.3) indicates a significant difference in favor of the experimental classes as compared with the comparison group ($p < .05$). However, the significance of these differences disappears when the pre-test 100-item Gumpgookies scores are used in a covariance analysis, although the post-test means are still in favor of the experimental groups (Table 7.4).

TABLE 7.1

Analysis of Variance of Stanford-Binet Pre- and Post-Test
Scores for Three Motivation Classes Combined (N = 59) and
Three Comparison Classes Combined (N = 41)

Groups	Means	
	Pre-Test	Post-Test
Motivation	87.95	93.19
Comparison	83.54	90.12

Sources of Variance

Source	Mean Squares	D.F.	F Ratio	p
Total	175.72	199		
Between Groups	275.26	99		
Groups	677.00	1	2.50	< .11
Error (Groups)	271.16	98		
Within Groups	77.17	100		
Trials	1677.00	1	27.30	< .001
Groups by Trials	20.00	1	.33	< .58
Error (Trials)	61.42	98		

TABLE 7.2

Analysis of Variance of 100-Item Gumpgookies
 Pre- and Post-Test Scores for Three Motivation Classes
 Combined (N = 59) and Three Comparison Classes Combined (N = 41)

Groups	Means	
	Pre-Test	Post-Test
Motivation	57.00	64.17
Comparison	56.78	60.95

Sources of Variance

Source	Mean Squares	D.F.	F Ratio	P
Total	83.36	199		
Between Groups	106.55	99		
Groups	142.69	1	1.34	< .25
Error (Groups)	106.18	98		
Within Groups	60.41	100		
Trials	1764. 00	1	41.45	< .001
Groups by Trials	109.13	1	2.57	< .11
Error (Trials)	42.53	98		

TABLE 7.3

Analysis of Variance of 75-Item,
Randomized Form of Gumpgookies for Three Classes

	Mean	N
Motivation 1	55.84	19
Motivation 2	53.40	15
Comparison	48.36	11

Sources of Variance

Source	Mean Squares	D.F.	F Ratio	P
Total	66.35	44		
Groups	195.25	2	3.24	< .05
Error	60.21	42		

TABLE 7.4

Analysis of Covariance of Scores
for the 75-Item, Randomized Form of Gumpgookies
With the Pre-Test 100-Item, Left-Right Form of
Gumpgookies as Covariate for Three Classes

Group	Mean	Adjusted Mean	Adjusted S.E.	N
Motivation 1	55.84	54.69	1.79	19
Motivation 2	53.40	53.49	1.93	15
Comparison	48.36	50.22	2.41	11

CHAPTER 8

SUMMARY AND PROPOSED FUTURE RESEARCH

This report has presented results to date of research over a five-year period, first, on measurement of motivation of young children to achieve in school. After consideration and tryout of numerous ideas and techniques, efforts were concentrated on development of the measuring instrument called Gumpgookies. Each item in this test consists of two imaginary figures called gumpgookies that the examiner describes. The child chooses the one with which he identifies. Three forms of the test have resulted from the preliminary analyses: (a) a 75-item form to be individually administered to preschool children, (b) one 100-item group form for non-reading elementary school children, and (c) a second 100-item group form for elementary school children who can read.

The test had its roots in a theoretical framework, which conceives of five constituents of motivation to achieve: (a) an affective component, expressed as positive affect from achievement; (b) a conceptual component, whereby the individual sees himself as an achiever; (c) a purposive component, enabling the individual to establish and respond to future goals; (d) a cognitive component, by means of which the instrumental steps necessary to attain goals are known; and (e) an ethical component, through which the individual can evaluate his own performance.

With the first format used for the test, techniques of factor analysis and a type of cluster analysis were applied in order to judge the adequacy of the underlying theory. It was soon discovered that the factors were partially determined by one or both of two extraneous influences: the position of the correct alternative on the item page--right or left--and

the order in which the two alternatives were presented to the child. With the original format, the effects of these two influences could not be separated, because the alternatives were read by the examiner from a left-hand position to a right-hand position. An additional extraneous influence affecting some factors seemed to be the position of an item in the test--toward the beginning or toward the end. The body of the report describes the various methods used in an attempt to shed light on these non-substantive influences.

The format was revised so that the illustrations appeared in various positions on the item pages, and the order in which the illustrations are described was randomized. At the same time, many of the items for the individual form for use with younger children were revised to reduce the cognitive and verbal difficulty.

The individual form was administered to 10 ethnic-cultural groups of preschool children, 1,607 in all, that were geographically scattered over the United States. The appropriate group form was administered to children in the first, second, and fourth grades in Hawaii, a total of 668.

Separate factor analyses (always to a six-factor solution and occasionally to solutions with other numbers of factors) were performed for each of the 10 ethnic-cultural groups and for each of the three grade groups, as well as for the 10 ethnic-cultural groups combined, grades 1 and 2 combined, and grades 1, 2, and 4 combined. The correlation coefficients among the loadings of the 60 factors for the 10 groups of preschool children were themselves factored. Likewise, the correlation coefficients for the 18 factors for grades 1, 2, and 4 were factored. By reference to the original factor solutions, items were assigned in each

case to one of six factors that have been referred to as initial super-factors. This technique involved a degree of subjectivity in assignment of items, and later efforts were addressed to the improved super-factors that resulted from factorizations of data for the large, combined groups.

The two improved super-factor solutions, which were not expected to be identical, could be compared because the 75 items in the form for younger children had been selected from the 100 used for the older groups. Some items, however, had been revised in wording, and the answer positions and order of presentation of alternatives were not necessarily the same.

For some of the super-factors, as well as for some factors based on the several subgroups, there is again evidence of the consequences of the three types of extraneous influences mentioned above, which seem to be "characteristics of the organism." The effects of answer position seem to be more prominent for the younger children, while the influence of primacy versus recency appears to be more in evidence for the older children. The position of items in the test also apparently affects certain factors for the younger children to a greater degree than for the older.

Despite these troublesome non-substantive influences that add to the difficulty of interpreting separate factors, the total score, when considered for a group of subjects, should not be systematically affected by them unless all of the subjects have the same answer-position, item-position, and primacy versus recency tendencies. Inspection of answer sheets reveals no such similarities. However, there seems to be no escape from the fact that total scores for individual children may be distorted by their idiosyncratic proclivities to be affected by irrelevant tendencies.

The super-factors based on the two analyses for the combined groups of subjects were interpreted. Some factors were more strongly identified than others, perhaps at least in part because production of items tapping certain components is relatively easy. For the elementary school children, interpretations for five factors that are consonant with the theoretical constituents were offered. A sixth, rather weak factor, not provided for in the theory, was very gingerly identified as self-reliance or dependability. Tentative interpretations for five of the six factors based on the combined preschool groups were made, and the relations with the six factors based on the older children were explored. If, as much evidence indicates, motivation represents learned behavior, factors will almost certainly differ for different age groups.

Other evidence relevant to the reliability and validity of the total score on Gumpgookies is collected in Chapter 5 of the report. It has long been clear that internal consistency estimates, based on KR-20, are quite satisfactory, running in the low to mid .80's for the 75-item form and in the upper .80's or low .90's for the 100-item form. Evidence for test-retest reliability or comparable forms reliability is not yet available, although it was found that scores on a 100-item, non-randomized form correlated .66 with scores on a 75-item, randomized, and revised form given two or three weeks later ($N = 44$). For the age ranges concerned, the correlation of Gumpgookies with age was low and positive, and its relation with IQ was of the order of .20 to .35. For one sample of 135 four-year-olds, its correlation with the Preschool Inventory was

.31, higher than the correlation of .24 with the Binet for that sample.* For a group of second-grade children, its correlation with the Children's Self-Concept Index, a test used in the Westinghouse Learning Corporation follow-up study of Head Start children based on one by Joseph H. Meyerowitz, was .43. This can be regarded as a relatively high correlation because the range of scores on the latter test was very small and because its content appears to deal principally with the Gumpgookies factor interpreted as self-confidence.

Chapter 5 also reports positive relations of Gumpgookies with teachers' rankings or ratings of motivation based upon several procedures and indicates that the test does discriminate between children rated as high and those rated as low in motivation.

Based upon the conviction that motivated behavior is learned, concurrent with the development and refinement of a measuring instrument were exploratory attempts to teach motivation in preschool classes, a second major focus of the studies reported herein. A motivation curriculum was developed and tried out with three Head Start classes, their scores on several pre- and post-tests being compared with those for three classes not exposed to the curriculum. The findings were considered of sufficient promise to warrant further efforts to teach motivation with a revised curriculum during the school year 1969-70.

*For that sample, in a factor analysis done after Chapter 6 was written, Gumpgookies had its highest loading on a factor also most heavily loaded for percentage of work responses on the Binet, motivation rankings, the Preschool Inventory, and the Animal-House subtest of the WPPSI. Perhaps this is a motivated work factor.

Avenues of Further Research

Further studies of Gumpgookies based upon data already available.

Although a large amount of data on Gumpgookies has been collected and much of it has been analyzed, more intensive study of the data is planned.

This work will be geared to the following problems:

1. Interpretation of the factors for each of the 13 subgroups and more detailed comparisons among these factors and of these factors with super-factors. Although many such comparisons have been made by systematic inspectional techniques as computer outputs became available, more rigorous comparisons by means of correlating each set of loadings with every other set will be more definitive.
2. Determination of the mean item difficulty level and the dispersion of item difficulty indices for each factor for each group.
3. Separate factor analyses for the two sexes for each of the subgroups.
4. Research on the relative importance of extraneous influences on the factors for the various subgroups.
5. Study of the individual items--their difficulty, item-test correlation, and factorial composition for each group--followed by item revision or rejection, if indicated.
6. Investigation of changes in item difficulty and item-test correlation with age, based upon different groups of children of different ages.

7. Study of changes in item difficulty and item-test correlation for the same sample of children who have been in a Head Start program from pre-test to post-test, over an interval of five to six months.
8. Estimation of the internal consistency of scores on super-factors.
9. Comparison of scores on the two sets of improved super-factors for the relevant subgroups.
10. A series of similar item analyses and factor analyses for the data on the first 100-item form used as a pre-test and a selected 55 items used as a post-test in the 1968-69 national evaluation study of Head Start. The total number of subjects for the 12 Evaluation and Research Centers was 1,742 for the pre-test and 1,624 for the post-test. Such studies, based upon subjects of different ethnic characteristics, from different parts of the United States, and differing with respect to urban versus rural environments, will also add valuable results for comparison with findings for the other groups already studied.

Further studies of Gumpgookies requiring additional data. Some of the problems that should be investigated will require additional data. These problems include the following:

1. A more dependable estimate of test-retest reliability than is now available will be made, based upon an interval of perhaps two to four weeks.
2. Eventually, if two comparable forms of the test can be developed, an estimate of comparable forms reliability will be obtained, along with such estimates for each major factor.

3. Further work on the relation of the test with external criteria will be done. Several of the criterion measures that have been tried in a limited way are promising. They, as well as new measures, warrant systematic study in relation to the test.
4. In addition to teacher ratings of general motivation, ratings for each of several factors may be useful as criteria for the respective factors. Such ratings have not yet been tried and teachers may be resistive, but the approach will be considered.
5. Criterion measures based upon records and/or products of actual behavior will be intensively explored. Such measures might be based upon performance in situational tests designed to present opportunities for motivated behavior to manifest itself with respect to one or more or even the totality of the factors. Once reliable criteria are available, further validation studies will be feasible.
6. The current forms of the test will be revised in the light of all data collected, and the reliability and validity of the revised forms will be investigated.
7. Since production of items tapping some factors has been easier than for others, more new items will be constructed for some of the less clearly identified factors so that eventually there can be scores for factors, not just total scores, that are reasonably reliable.
8. More nearly comparable samples of subjects in different ethnic, cultural, and socioeconomic groups will be needed before fully defensible comparisons among groups can be made.

9. In relation to the development of motivation, possible interactions among ethnicity, socioeconomic level, rural versus urban background, sex, and possibly age and other variables warrant study.
10. Once significant differences among subgroups are identified, the etiology of such differences--as, e.g., with respect to child-rearing practices; parental attitudes, aspirations, and expectations with respect to the child; the parents' own motivation to achieve; teacher attitudes; preschool and elementary school curricula; and a host of other factors--will be studied.
11. Methodological solutions are at hand that will further the clarification of substantive motivational factors, free of the extraneous effects of the vexatious set factors that have hampered interpretations of motivational factors. One such solution, for a test of the length regarded as essential to tap several factors, is suggested by the work of Sheila C. Feld and Judith Lewis (1969) on the Test Anxiety Scale for Children (Sarason, et al., 1960). They developed special forms of the test, each including some items in reversed form as well as in the original form, in an effort to separate the effects of an acquiescence factor from the content factors in the test. They then argued, quite soundly, that a factor on which the reversed item tended to have loadings about equal but of opposite sign to that of the original item was a genuine content factor. With the Gumpgookies test, however--and especially in view of the number of factors that are hypothesized and that hence require a relatively large number of items--another approach

is preferred. It will not require giving the same items to the same group of children in one of several possible reversed forms. Rather, it will necessitate that different forms of the test be given to different groups of subjects. All forms of the test will contain identical items insofar as content is concerned; but answer positions, item positions in the test, and order of presentation of alternatives for each item will be systematically varied. The number of different forms will be dependent in large part upon financial and personnel resources available for the study. In any case, all results, item by item, will be combined, and perhaps factor analyses based on subsamples representing different populations can be made as well. The results, it is expected, will confirm that there are substantive motivational factors, independent of extraneous set factors.

12. Following the foregoing research that will serve to separate the influence of motivational content factors from extraneous set factors, more work will be done to develop criteria for establishing the extent to which set factors may be distorting or invalidating scores of individual children.

Further efforts toward curriculum development for preschool children.

Accompanying work aimed at improving the measuring instrument with respect to reliability and validity and at understanding the basic constituents of motivation will be more intensified efforts in the direction of teaching motivation. The work that has been reported is correctly to be regarded as in a very early stage. Much more needs to be known about

1. the extent to which effects of training in motivation are retained, as might be investigated by means of follow-up

studies on groups such as those in the 1968-69 experimental classes and later ones.

2. kinds of activities conducive to development of particular factors of motivation--and for what kinds of children at what ages.
3. ways of introducing the activities and related materials in a class setting, taking into account optimal sequencing.
4. characteristics of teachers that are associated with their abilities to inculcate motivation in the children they teach.
5. techniques and materials to be used to train teachers in improved abilities to teach the various facets of motivation.
6. methods and materials for training parents to reinforce efforts of teachers to induce the pervasive habits and attitudes subsumed under the term motivation.

Eventual extension of techniques for teaching motivation to pre-preschoolers and post-preschoolers. Children do not suddenly begin to be motivated when they reach preschool age, nor should they cease learning to be motivated when they leave preschool. Undoubtedly certain types of environments and experiences tend to foster one or more components of motivation at very young ages. If such factors can be identified, they almost certainly will not be interpretable as identical with those for preschool children or for elementary school children. But perhaps even for children as young as two or three years of age broader or more general factors can be conceptualized, though their accurate measurement may not be possible, and serve as a springboard for attempting to stimulate and reinforce motivated behavior. Such factors, just by way of illustration, might be energetic behavior; self-confidence, which, to be sure, shows up for the older two levels; responsiveness to immediate

goals, in contrast to more remote goals; responsiveness to the influence of adults or to that of other persons; attentiveness to things in the environment; curiosity; and so on.

If a group of experts in the learning of children of from two to three years of age could agree upon potentially fruitful focuses of training efforts, it should be possible to devise curricular materials and teaching procedures that would enhance behavior in the predicated directions. Clearly such efforts would have to intimately involve parents or whatever persons are responsible for care of the young child.

At the other end of the age spectrum beyond preschool, efforts to teach motivation should be sustained. Indeed, if the ages six to nine are especially critical in the evaluation of motivation to achieve (Veroff, 1969), then this age period may be optimal for efforts to teach motivation. Some of the general ideas and methods here developed for preschool children may continue to be effective for older children, but additional innovative efforts will be needed.

Appendix

"Norms" for the 55-Item Gumpgookies, Based Upon Pre-test Scores of 1485 Children in the 1968-69 Head Start National Evaluation Sample

University of Hawaii
Head Start Research Center

The nature of these "norms" and the procedure by which they were developed calls for some explanation. First, scores on the Gumpgookies pre-test, which contained 100 items, were recomputed for the 55 items that comprised the post-test. The subjects were then segregated into age groups, with a one-month interval. Data from all Head Start Evaluation and Research Centers were supplied to the Hawaii Center by Dr. Lois-ellin Datta. The means and standard deviations for these age groups are shown in an accompanying table:

Means and Standard Deviations on the 55 Gumpgookies
Pre-test Items that Comprised the Post-test, by One-Month Age Intervals

<u>Age</u>	<u>Mean</u>	<u>Standard Deviation</u>	<u>N</u>	<u>Age</u>	<u>Mean</u>	<u>Standard Deviation</u>	<u>N</u>
69	41.1	7.4	27	56	35.8	9.3	99
68	41.4	7.5	23	55	36.4	8.1	78
67	41.5	8.4	35	54	37.3	8.4	80
66	40.1	6.8	39	53	36.7	9.1	71
65	39.9	8.0	31	52	36.2	8.1	109
64	42.3	7.6	36	51	33.9	8.6	96
63	39.7	7.6	49	50	33.9	8.1	80
62	41.6	8.1	54	49	37.5	8.8	102
61	38.4	6.9	39	48	33.2	9.9	74
60	39.8	9.0	32	47	35.05	6.3	58
59	38.0	7.0	56	46	33.6	8.0	45
58	36.2	9.3	61	45	34.1	7.2	18
57	37.3	8.1	72	44	34.2	6.7	21

It will be noted that, although there is a trend for the scores to increase with age, the curve of mean score against age fluctuates. For example, the mean for children 49 months of age exceeds those for children 50 and 51 months of age; and the mean for the 62-month age group exceeds the means for ages 65 through 69 months. This finding, although not unanticipated, indicates that it would not make sense to try to develop "national" norms, such as transmuted standard scores, using for each monthly age group its mean and standard deviation. The reason is simply that a particular score for one age group might have a lower transmuted score than would the same score for a higher age group. Nor would the irregularity disappear entirely if a larger age interval, say two, three, or four months, were to be used, and a still larger interval would defeat the purpose of the norms.

The procedure adopted was to "fit" a straight line to the means for successive age groups, to extrapolate it downward to age 43 months and upward to 79 months, then to read from the graph the mean corresponding to each monthly age interval. The standard deviations are not constant but reveal no particular trend with increasing age. The unweighted average is 7.7. To simplify the arithmetic, the following formula was applied:

$$Z = \frac{15 (X - \bar{X})}{7.5} + 100,$$

where Z is the transmuted score and X the raw score. This formula will result in scores with a predicted mean of 100 and a standard deviation of roughly 15. When the formula was applied, rounding resulted in the same Z-scores for particular adjacent age groups. Hence the ages in questions were combined in the resulting table of norms.

It is recognized that the data are not ideal mainly because the different age groups from different parts of the country do not represent large samples drawn at random from a nation-wide pool of Head Start children. Nevertheless, it is hoped that they will be serviceable in interpreting change from pre-test to post-test, as against trying to interpret raw-score changes without an age base. Each Center is of course free to make whatever use of the table of norms it desires, or no use.

Z-Score "Norms" for the 55-Item Gumpgookies, Based Upon Pre-test Scores of 1485 Children in the 1968-69 Head Start National Evaluation Sample (See Text)

	Age in Months													
	44-		48-		51-		54-		58-					
	43	45	46	47	49	50	52	53	55	56	57	59	60	61
55	144	143	142	141	140	139	138	137	136	135	134	133	132	131
54	142	141	140	139	138	137	136	135	134	133	132	131	130	129
53	140	139	138	137	136	135	134	133	132	131	130	129	128	127
52	138	137	136	135	134	133	132	131	130	129	128	127	126	125
51	136	135	134	133	132	131	130	129	128	127	126	125	124	123
50	134	133	132	131	130	129	128	127	126	125	124	123	122	121
49	132	131	130	129	128	127	126	125	124	123	122	121	120	119
48	130	129	128	127	126	125	124	123	122	121	120	119	118	117
47	128	127	126	125	124	123	122	121	120	119	118	117	116	115
46	126	125	124	123	122	121	120	119	118	117	116	115	114	113
45	124	123	122	121	120	119	118	117	116	115	114	113	112	111
44	122	121	120	119	118	117	116	115	114	113	112	111	110	109
43	120	119	118	117	116	115	114	113	112	111	110	109	108	107
42	118	117	116	115	114	113	112	111	110	109	108	107	106	105
41	116	115	114	113	112	111	110	109	108	107	106	105	104	103
40	114	113	112	111	110	109	108	107	106	105	104	103	102	101
39	112	111	110	109	108	107	106	105	104	103	102	101	100	99
38	110	109	108	107	106	105	104	103	102	101	100	99	98	97
37	108	107	106	105	104	103	102	101	100	99	98	97	96	95
36	106	105	104	103	102	101	100	99	98	97	96	95	94	93
35	104	103	102	101	100	99	98	97	96	95	94	93	92	91
34	102	101	100	99	98	97	96	95	94	93	92	91	90	89
33	100	99	98	97	96	95	94	93	92	91	90	89	88	87
32	98	97	96	95	94	93	92	91	90	89	88	87	86	85
31	96	95	94	93	92	91	90	89	88	87	86	85	84	83
30	94	93	92	91	90	89	88	87	86	85	84	83	82	81
29	92	91	90	89	88	87	86	85	84	83	82	81	80	79
28	90	89	88	87	86	85	84	83	82	81	80	79	78	77
27	88	87	86	85	84	83	82	81	80	79	78	77	76	75
26	86	85	84	83	82	81	80	79	78	77	76	75	74	73
25	84	83	82	81	80	79	78	77	76	75	74	73	72	71
24	82	81	80	79	78	77	76	75	74	73	72	71	70	69
23	80	79	78	77	76	75	74	73	72	71	70	69	68	67
22	78	77	76	75	74	73	72	71	70	69	68	67	66	65
21	76	75	74	73	72	71	70	69	68	67	66	65	64	63
20	74	73	72	71	70	69	68	67	66	65	64	63	62	61
19	72	71	70	69	68	67	66	65	64	63	62	61	60	59
18	70	69	68	67	66	65	64	63	62	61	60	59	58	57
17	68	67	66	65	64	63	62	61	60	59	58	57	56	55
16	66	65	64	63	62	61	60	59	58	57	56	55	54	53
15	64	63	62	61	60	59	58	57	56	55	54	53	52	51
14	62	61	60	59	58	57	56	55	54	53	52	51	50	49
13	60	59	58	57	56	55	54	53	52	51	50	49	48	47
12	58	57	56	55	54	53	52	51	50	49	48	47	46	45
11	56	55	54	53	52	51	50	49	48	47	46	45	44	43
10	54	53	52	51	50	49	48	47	46	45	44	43	42	41

Z-Score "Norms" for 55-Item Gumpgookies, Continued

Age in Months

Raw Score

	62- 63	64	65- 66	67	68	69- 70	71	72	73- 74	75	76- 77	78	79
55	130	129	128	127	126	125	124	123	122	121	120	119	118
54	128	127	126	125	124	123	122	121	120	119	118	117	116
53	126	125	124	123	122	121	120	119	118	117	116	115	114
52	124	123	122	121	120	119	118	117	116	115	114	113	112
51	122	121	120	119	118	117	116	115	114	113	112	111	110
50	120	119	118	117	116	115	114	113	112	111	110	109	108
49	118	117	116	115	114	113	112	111	110	109	108	107	106
48	116	115	114	113	112	111	110	109	108	107	106	105	104
47	114	113	112	111	110	109	108	107	106	105	104	103	102
46	112	111	110	109	108	107	106	105	104	103	102	101	100
45	110	109	108	107	106	105	104	103	102	101	100	99	98
44	108	107	106	105	104	103	102	101	100	99	98	97	96
43	106	105	104	103	102	101	100	99	98	97	96	95	94
42	104	103	102	101	100	99	98	97	96	95	94	93	92
41	102	101	100	99	98	97	96	95	94	93	92	91	90
40	100	99	98	97	96	95	94	93	92	91	90	89	88
39	98	97	96	95	94	93	92	91	90	89	88	87	86
38	96	95	94	93	92	91	90	89	88	87	86	85	84
37	94	93	92	91	90	89	88	87	86	85	84	83	82
36	92	91	90	89	88	87	86	85	84	83	82	81	80
35	90	89	88	87	86	85	84	83	82	81	80	79	78
34	88	87	86	85	84	83	82	81	80	79	78	77	76
33	86	85	84	83	82	81	80	79	78	77	76	75	74
32	84	83	82	81	80	79	78	77	76	75	74	73	72
31	82	81	80	79	78	77	76	75	74	73	72	71	70
30	80	79	78	77	76	75	74	73	72	71	70	69	68
29	78	77	76	75	74	73	72	71	70	69	68	67	66
28	76	75	74	73	72	71	70	69	68	67	66	65	64
27	74	73	72	71	70	69	68	67	66	65	64	63	62
26	72	71	70	69	68	67	66	65	64	63	62	61	60
25	70	69	68	67	66	65	64	63	62	61	60	59	58
24	68	67	66	65	64	63	62	61	60	59	58	57	56
23	66	65	64	63	62	61	60	59	58	57	56	55	54
22	64	63	62	61	60	59	58	57	56	55	54	53	52
21	62	61	60	59	58	57	56	55	54	53	52	51	50
20	60	59	58	57	56	55	54	53	52	51	50	49	48
19	58	57	56	55	54	53	52	51	50	49	48	47	46
18	56	55	54	53	52	51	50	49	48	47	46	45	44
17	54	53	52	51	50	49	48	47	46	45	44	43	42
16	52	51	50	49	48	47	46	45	44	43	42	41	40
15	50	49	48	47	46	45	44	43	42	41	40	39	38
14	48	47	46	45	44	43	42	41	40	39	38	37	36
13	46	45	44	43	42	41	40	39	38	37	36	35	34
12	44	43	42	41	40	39	38	37	36	35	34	33	32
11	42	41	40	39	38	37	36	35	34	33	32	31	30
10	40	39	38	37	36	35	34	33	32	31	30	29	28

REFERENCES

- Adkins, D. C. & Ballif, B. L. Exploration of motivation to achieve in preschool children. University of Hawaii Head Start Evaluation and Research Center Annual Report, 1966-67. Submitted to the Institute for Educational Development for the Research and Evaluation Division, Project Head Start, Office of Economic Opportunity.
- Adkins, D. C. & Ballif, B. L. Measurement of motivation to achieve in preschool children. Final Report on Contract No. OEO 4218. Submitted to the Office of Economic Opportunity, 1968.
- Anastasi, A. Psychological testing. New York: Macmillan, 1954.
- Argyle, M. & Robinson, P. Two origins of achievement motivation. British Journal of Social and Clinical Psychology, 1962, 1, 107-120.
- Atkinson, J. W. Explorations using imaginative thought to assess the strength of human motives. In M. R. Jones (Ed.), The Nebraska symposium on motivation. Lincoln: University of Nebraska Press, 1954.
- Atkinson, J. W. Motivational determinants of risk-taking behavior. Psychological Review, 1957, 64, 359-372.
- Atkinson, J. W. The mainsprings of achievement-oriented activity. In J. D. Krumboltz (Ed.), Learning and the educational process. Chicago: Rand McNally, 1965. Pp. 25-66.
- Bandura, A. Social learning through imitation. In M. R. Jones (Ed.), Nebraska symposium on motivation 1962. Lincoln: University of Nebraska Press, 1962. Pp. 211-268.
- Bandura, A., Grusec, J. E., & Menlove, F. L. Some social determinants of self-monitoring reinforcement systems. Journal of Personality and Social Psychology, 1967, 5, 449-455.
- Bandura, A. & Kupers, C. J. Transmission of self-reinforcement through modeling. Journal of Abnormal and Social Psychology, 1964, 69, 1-9.
- Bandura, A. & Whalen, C. K. The influence of antecedent reinforcement and divergent modeling cues on patterns of self-reward. Journal of Personality and Social Psychology, 1966, 3, 373-382.
- Barlow, J. A. & Burt, C. The identification of factors from different experiments. British Journal of Statistical Psychology, 1954, 7, 52-56.
- Berkowitz, H. & Zigler, E. Effects of preliminary positive and negative interactions and delay conditions on children's responsiveness to social reinforcement. Journal of Personality and Social Psychology, 1965, 2, 500-505.

- Bloxom, B. Factorial rotation to simple structure and maximum similarity. Psychometrika, 1968, 33, 237-247.
- Brim, O. G., Jr. Personality development as role learning. In I. Iscoe & H. W. Stevenson (Eds.), Personality development in children. Austin: University of Texas Press, 1960. Pp. 127-159.
- Brookover, W. B., Erickson, E. L., & Joiner, L. M. Self-concept and student role achievement. Paper presented at American Educational Research Association Meetings, Los Angeles, February 1969.
- Bryan, J. F. & Locke, E. A. Goal setting as a means of increasing motivation. Journal of Applied Psychology, 1967, 51, 274-277.
- Burt, C. The factors of the mind. New York: Macmillan, 1941.
- Carroll, J. B. The effect of difficulty and chance success on correlations between items or between tests. Psychometrika, 1945, 10, 1-20.
- Cattell, R. B., Balcar, K. R., Horn, J. L., & Nesselroade, J. R. Factor matching procedures: An improvement of the s index; with tables. Educational and Psychological Measurement, 1969, 29, 781-792.
- Cofer, C. N. & Appley, M. H. Motivation: Theory and research. New York: Wiley, 1964.
- Combs, A. W. & Soper, D. W. The relationship of child perceptions to achievement and behavior in the early school years. Gainesville, Florida: University of Florida, 1963.
- Congor, A. J. Dimensional analyses of binary data: An appraisal of the MMPI. Project on Techniques for Investigation of Structure of Individual Differences in Psychological Phenomena. Department of Psychology, University of Illinois, Urbana, Illinois, 1968.
- Coopersmith, S. Implications of studies on self-esteem for educational research and practice. Paper presented at American Educational Research Association Meetings, Los Angeles, February 1969.
- Digman, J. M. A test of a multiple-factor model of child personality. Honolulu: University of Hawaii, 1965.
- Dingman, H. F. The relation between coefficients of correlation and difficulty factors. British Journal of Statistical Psychology, 1958, 11, 13-17.
- Doob, L. W. The behavior of attitudes. Psychological Review, 1947, 54, 135-156.
- Dreyer, A. S. & Haupt, D. Self-evaluation in young children. Journal of Genetic Psychology, 1966, 108, 185-197.

- Feld, S. C. & Lewis, J. The assessment of achievement anxieties in children. In C. P. Smith (Ed.), Achievement-related motives in children. New York: Russell Sage Foundation, 1969. Pp. 151-199.
- Ferguson, G. A. The factorial interpretation of test difficulty. Psychometrika, 1941, 6, 323-330.
- Flanders, J. P. A review of research on imitative behavior. Psychological Bulletin, 1968, 69, 316-337.
- Gilbert, D. C. The young child's awareness of affect. Child Development, 1969, 40, 629-640.
- Gordon, E. W. & Wilkerson, D. A. Compensatory education for the disadvantaged. New York: College Entrance Examination Board, 1968.
- Gourlay, N. Difficulty factors arising from the use of tetrachoric correlations in factor analysis. British Journal of Psychology, 1951, Stat. Section 4, 65-76.
- Gray, S. W. & Klaus, R. A. An experimental preschool program for culturally deprived children. Child Development, 1965, 36, 887-898.
- Guilford, J. P. The difficulty of a test and its factor composition. Psychometrika, 1941, 6, 67-78.
- Haber, R. N. Discrepancy from adaptation level as a source of affect. Journal of Experimental Psychology, 1958, 56, 370-375.
- Hardy, K. R. An appetitional theory of sexual motivation. Psychological Review, 1964, 71, 1-18.
- Hebb, D. O. The organization of behavior. New York: Wiley, 1949.
- Helson, H. Adaptation level as frame of reference for prediction of psychological data. American Journal of Psychology, 1947, 60, 1-29.
- Henderson, R. W. Influences on cognitive and motivational conditions for acculturation. Paper read at American Educational Research Association Meetings, New York, February 1967.
- Herbert, E. W., Gelfand, D. M., & Hartmann, D. P. Imitation and self-esteem as determinants of self-critical behavior. Child Development, 1969, 40, 421-430.
- Hilgard, E. R. & Bower, G. H. Theories of learning. New York: Appleton-Century-Crofts, 1968.
- Hunt, J. McV. The role of experience in the development of intelligence. Invited address for Psi Chi, at Midwestern Psychological Association Meetings, Chicago, May 1967.

- Kagan, J. Motivational and attitudinal factors in receptivity to learning. In J. Bruner (Ed.), Learning about learning. Washington, D. C.: U. S. Government, 1966.
- Kaiser, H. F. The application of electronic computers to factor analysis. Educational and Psychological Measurement, 1960, 20, 141-151.
- Katz, I. The socialization of academic motivation in minority group children. In D. Levine (Ed.), Nebraska symposium on motivation 1967. Lincoln: University of Nebraska Press, 1967. Pp. 133-191.
- Kohlberg, L. Development of moral character and ideology. In M. L. Hoffman (Ed.), Review of child development research, Vol. 1. New York: Russell Sage Foundation, 1964.
- Liebert, R. M. & Allen, M. K. The effects of rule structure and reward magnitude on the acquisition and adoption of self-reward criteria. Psychological Reports, 1967, 21, 445-452.
- Liebert, R. M. & Ora, J. P., Jr. Children's adoption of self-reward patterns: Incentive level and method of transmission. Child Development, 1968, 39, 537-544.
- Malzman, I. Thinking: From a behavioristic point of view. Psychological Review, 1955, 62, 275-286.
- Martine, J. G. Relationship between the self concept and differences in the strength and generality of achievement motivation. Journal of Personality and Social Psychology, 1956, 24, 364-375.
- Mattick, I. Adaptation of nursery school techniques to deprived children. Journal of the Academy of Child Psychiatry, 1965, 4, 670-700.
- McClelland, D. C. Personality. New York: Sloane, 1951.
- McClelland, D. C. Methods of measuring human motivation. In J. W. Atkinson (Ed.), Motives in fantasy, action, and society. New York: D. Van Nostrand, 1958. Pp. 7-42.
- McClelland, D. C. Toward a theory of motive acquisition. American Psychologist, 1965, 20, 321-333.
- McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. The achievement motive. New York: Appleton-Century-Crofts, 1953.
- McCoy, N. & Zigler, E. Social reinforcer effectiveness as a function of the relationship between child and adult. Journal of Abnormal and Social Psychology, 1965, 1, 604-612.
- Miller, N. E. Comments on theoretical models illustrated by the development of a theory of conflict behavior. Journal of Personality and Social Psychology, 1951, 20, 82-106.

Mischel, W. & Liebert, R. M. Effects of discrepancies between observed and imposed reward criteria on their acquisition and transmission. Journal of Personality and Social Psychology, 1966, 3, 45-53.

Ofstad, N. S. The transmission of self-reinforcement patterns through imitation of sex-role appropriate behavior. Unpublished doctoral dissertation, University of Utah, 1960.

Olds, J. Physiological mechanisms of reward. In M. R. Jones (Ed.), The Nebraska symposium on motivation. Lincoln: University of Nebraska Press, 1955.

Pederson, D. M. Ego strength and discrepancy between conscious and unconscious self-concepts. Perceptual and Motor Skills, 1965, 20, 691-692.

Rasey, M. I. & Menge, J. W. What we learn from children. New York: Harper & Brothers, 1956.

Raynor, J. O. The functional significance of future goals. Ann Arbor: The University of Michigan, 1967.

Sarason, S. B. et al. Anxiety in elementary school children: A report of research. New York: Wiley, 1960.

Savitsky, C. Reaching the disadvantaged. In E. P. Torrance & R. D. Strom (Eds.), Mental health and achievement. New York: Wiley, 1965.

Sears, P. Attitudinal and affective factors in children's approaches to problem-solving. In J. Bruner (Ed.), Learning about learning. Washington, D. C.: U. S. Government, 1966. Pp. 28-33.

Smith, L. Significant differences between high ability achieving and non-achieving college freshmen as revealed by interview data. Journal of Educational Research, 1965, 59, 10-12.

Staats, A. W., Staats, U. K., & Heard, W. G. Attitude development and ratio of reinforcement. Sociometry, 1960, 23, 338-350.

Todd, F. G., Terrell, C., & Frank, C. E. Differences between normal and underachievers of superior ability. Journal of Applied Psychology, 1962, 46, 183-190.

Torrance, E. P. Motivating children with school problems. In E. P. Torrance & R. D. Strom (Eds.), Mental health and achievement. New York: Wiley, 1965.

Veroff, J. Social comparison and the development of achievement motivation. In C. P. Smith (Ed.), Achievement-related motives in children. New York: Russell Sage Foundation, 1969. Pp. 46-101.

Wrigley, C. & Neuhaus, J. E. The matching of two sets of factors. (Abstract of a paper presented at the 63rd annual convention of the American Psychological Association, September 1-7, 1955) American Psychologist, 1955, 10, 418-449.

Young, P. T. The role of hedonic processes in motivation. In M. R. Jones (Ed.), The Nebraska symposium on motivation. Lincoln: University of Nebraska Press, 1955.

Young, P. T. The role of affective processes in learning and motivation. Psychological Review, 1959, 66, 104-125.

Zigler, E. & Butterfield, E. C. Motivational aspects of changes in IQ test performance of culturally deprived nursery school children. Child Development, 1968, 39, 1-14.